MetAegis: Defend Against Physical-layer Wireless Sensing Leakage via Metasurface Obfuscation

JINJIANG YANG*, Northwest University, China
CHENGHAN JIANG*, Northwest University, Shaanxi Key Laboratory of Passive Internet of Things and Neural Computing, Xi'an Key Laboratory of Advanced Computing and Software Security, China
JIE ZHANG, Xi'an University of Posts and Telecommunications, Northwest University, China
XINYI LI, Northwest University, Xi'an Advanced Battery-Free Sensing and Computing Technology International Science and Technology Cooperation Base, China
BOXUAN YANG, Northwest University, China
SHIBO ZHANG, Northwest University, China
XIANJIA MENG, Northwest University, China
ZHUZHU WANG, Northwest University, China
DINGYI FANG, Northwest University, Shaanxi International Joint Research Centre for the Battery-Free Internet of Things, China
ZHANYONG TANG†, Northwest University, Shaanxi Key Laboratory of Passive Internet of Things and Neural Computing, Xi'an Key Laboratory of Advanced Computing and Software Security, China

Wi-Fi-based non-intrusive sensing has emerged as a promising technology for monitoring human activities without the need for wearable devices, attracting significant interest from both industry and academia. However, the inherent broadcast nature of wireless signals renders user activity features vulnerable to eavesdropping by malicious entities, posing serious privacy risks. Existing privacy protection strategies predominantly rely on source-side solutions, such as hardware modifications to encrypt wireless channels. While partially effective, these approaches often necessitate costly and impractical hardware upgrades and, in some cases, fail to accommodate the needs of authorized users, thereby compromising their sensing performance in practical scenarios. To overcome these limitations, we propose MetAegis, an innovative channel-side obfuscation system that leverages a programmable metasurface for flexible and cost-effective deployment. MetAegis dynamically alters the wireless channel to achieve selective feature obfuscation across multiple directions, effectively shielding against malicious users while ensuring that authorized users experience no degradation in performance. This is accomplished through a feature obfuscation

*Both authors contributed equally to this research.
†Corresponding author.

Authors' Contact Information: Jinjiang Yang, Northwest University, China, yangjinjiang@stumail.nwu.edu.cn; Chenghan Jiang, Northwest University, Shaanxi Key Laboratory of Passive Internet of Things and Neural Computing, Xi'an Key Laboratory of Advanced Computing and Software Security, China, jianchanghan@stumail.nwu.edu.cn; Jie Zhang, Xi'an University of Posts and Telecommunications, Northwest University, China, jiezhang@xupt.edu.cn; Xinyi Li, Northwest University, Xi'an Advanced Battery-Free Sensing and Computing Technology International Science and Technology Cooperation Base, China, xinyili@nwu.edu.cn; Boxuan Yang, Northwest University, China, yangboxuan123@stumail.nwu.edu.cn; Shibo Zhang, Northwest University, China, zhangshibo@stumail.nwu.edu.cn; Xianjia Meng, Northwest University, China, xianjiam@nwu.edu.cn; Zhuzhu Wang, Northwest University, China, zzwang@nwu.edu.cn; Dingyi Fang, Northwest University, Shaanxi International Joint Research Centre for the Battery-Free Internet of Things, China, dyf@nwu.edu.cn; Zhanyong Tang, Northwest University, Shaanxi Key Laboratory of Passive Internet of Things and Neural Computing, Xi'an Key Laboratory of Advanced Computing and Software Security, China, zytang@nwu.edu.cn.

ACM 2474-9567/2026/3-ART22
https://doi.org/10.1145/3790113

technique driven by switching coding configurations on the metasurface, coupled with a feature recovery framework that enables authorized users to accurately reconstruct activity-related features from obfuscated signals. We implemented MetAegis using off-the-shelf Wi-Fi devices and conducted extensive evaluations across three real-world scenarios. Experimental results demonstrate that MetAegis reduces eavesdropper recognition accuracy to below 24% on average, while authorized users maintain average accuracy exceeding 88% through the recovery mechanism under the regular deployment. Remarkably, unlike traditional source-side solutions that require close integration with transmitters, MetAegis sustains its effectiveness even under challenging conditions, such as when the metasurface is positioned over 5 meters from the transmitter. These findings underscore the practicality and robustness of MetAegis in safeguarding privacy without sacrificing usability.

CCS Concepts: • **Security and privacy → Mobile and wireless security**.

Additional Key Words and Phrases: Wireless Sensing, Metasurface, Physical Layer Security

ACM Reference Format:
Jinjiang Yang, Chenghan Jiang, Jie Zhang, Xinyi Li, Boxuan Yang, Shibo Zhang, Xianjia Meng, Zhuzhu Wang, Dingyi Fang, and Zhanyong Tang. 2026. MetAegis: Defend Against Physical-layer Wireless Sensing Leakage via Metasurface Obfuscation. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 10, 1, Article 22 (March 2026), 28 pages. https://doi.org/10.1145/3790113

1 INTRODUCTION

The "sensorless" wireless sensing technology has become a hot topic in industry and academia in recent years. The key sensing principle is related to the disturbance caused by human activities [54, 69] or gestures [67, 68] on the wireless signal. Due to the wide deployment and low cost, Wi-Fi sensing stands out among a series of wireless sensing technologies such as acoustic [9, 52], visible light [26, 34], millimeter wave [33, 64], etc. Furthermore, Wi-Fi sensing technology claims many exciting applications in our daily lives, such as smart home and home health care [8, 61]. However, due to the open nature of the wireless medium, the Wi-Fi signal is at risk of being leaked to eavesdroppers, which causes the user's privacy to be illegally stolen. This privacy leakage includes, but is not limited to, the following potential consequences: eavesdroppers can use receivers placed near windows to understand user activity patterns or infer whether the user is at home (no signal fluctuations for a long time). According to statistics, 72% of burglaries occur when the homeowner is out [1]. In addition, eavesdroppers can even know the private activities of humans, which leads to serious risks of private leakage.

To address this potential threat, some excellent defense methods focus on implementing physical-layer encryption through hardware modifications or by deploying additional devices. Specifically, PhyCloak [44] pioneered the use of a carefully designed full-duplex repeater to modify the original signal. The same idea was also proposed in RF-Protect [49] by modifying the wireless signal, which injects fake human activity information designed to reflect signals to confuse eavesdroppers. However, the extra RF-chains will increase the defense burden. In recent years, a particularly effective countermeasure has been introduced by MimoCrypt [35], which employs a multiple-input multiple-output (MIMO) antenna array to both modulate and encrypt the Wi-Fi channel state information (CSI) signal. The approach successfully confuses eavesdroppers, and it allows legitimate users to maintain sensing capabilities through the use of a "key". But the requirement for the transceivers may limit the general promotion.

Facing these limitations, a critical question arises: Is there a plug-and-play obfuscation method without extra RF-chains to achieve Wi-Fi sensing protection? The rise of the metasurface has brought us a positive answer, which is a two-dimensional planar structure composed of a periodic arrangement of artificial atoms with special electromagnetic properties. Metasurface can dynamically reshape wireless signals in space. IRShield [51] has explored the feasibility by randomly switching coding configurations on the metasurface to confuse eavesdroppers' sensing results, but it ignores the original authorized sensing in the environment. To fill this gap, we proposed MetAegis[1], a plug-and-play Wi-Fi sensing obfuscation system. It releases the security task from the source side

[1]"MetAegis" is a combination of "Metasurface" and "Aegis", the shield in the hands of Zeus in ancient Greek mythology.

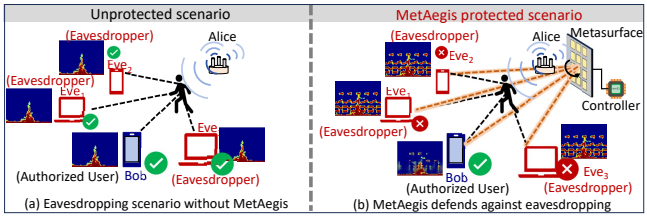


Fig. 1. Illustration of MetAegis protect case.

Table 1. Comparison of related works.

	Deploy type	Refuse Eve	Support Bob	Extra RF-chains
IRShield	Channel-side	✔	✖	No
MIMOCrypt	Source-side	✔	✔	No
WiShield	Source-side	✔	✔	No
PhyCloak	Channel-side	✔	✔	Yes
RF-Protect	Channel-side	✔	✔	Yes
Ours	Channel-side	✔	✔	No

to the channel side by deploying a smart metasurface near the Wi-Fi transmitter to serve as security "Aegis". However, MetAegisneeds to face three harsh facts:

C1) Limited Space For Obfuscation. Unlike active devices such as MIMO antenna arrays, metasurfaces can not actively transmit signals for obfuscation of the physical layer. Instead, they can only passively modulate incident electromagnetic waves due to the discrete states of their internal components, such as PIN diodes, and metasurfaces are limited to discretely regulating the amplitude or phase of the incident signal. Thus, how to obfuscate the feature under the limited space is a key challenge we need to solve. Although rapidly changing metasurface coding configurations can obfuscate the channel, it will make legitimate sensing tasks unsuitable. Consequently, achieving effective obfuscation within the constrained perturbation space presents a critical challenge.

C2) Limited Prior Knowledge of Eavesdroppers. In real-world physical layer security scenarios, the number and locations of potential eavesdroppers are often unknown in advance due to the lack of prior knowledge of eavesdroppers. It complicates the protection strategies against specific threats. Therefore, how to design a metasurface configuration strategy that can balance obfuscation and coverage performance to confuse potential eavesdroppers remains a practical challenge.

C3) Ensuring Authorized Sensing. The deployment of the metasurface introduces obfuscation strategies to counter unknown eavesdroppers. However, these strategies often degrade signal quality across the shared wireless medium, disrupting the CSI signal available to authorized users. As a result, authorized users receive obfuscated signals similar to those experienced by eavesdroppers, compromising their ability to perform reliable sensing seriously. In such a "lose-lose" fact, a critical challenge is how to enable authorized users to recover accurate information from these degraded, obfuscated signals, thereby maintaining sensing performance.

To address C1, we leverage the physical-layer obfuscation by treating the metasurface as a dynamic "secret key" that disturbs wireless channel states selectively. Specifically, the strategy is based on two interesting observations. We found that the different coding switching rates of the metasurface can generate disturbances of various basic frequencies. At the same time, by adjusting the different duration delays (i.e., duty cycle rate) of the metasurface codings, the high harmonic components of the disturbance frequency can be strategically controlled to achieve confusion of wireless signal features in multiple frequency components.

To solve C2, our basic idea is to use the beam scan technology to cover multiple directions. However, traditional beamforming beams are limited by a narrow "pencil beam" and can not achieve multiple directions. Therefore, a beam-broadening strategy that extends coverage is proposed. Specifically, we divide the directions as sub-areas and then make each beam cover about 30°. Then, utilize the optimization algorithm to search for effective beam steering that can "splice" a broad beam. Finally, we exploit the strategy proposed in C1 to switch the coding configurations in the coding set to achieve multiple directions protection.

To address C3, MetAegisintroduces an activity-related feature recovery strategy that bridges wireless sensing and image restoration tasks in Computer Vision (CV). Specifically, MetAegisleverages state-of-the-art image reconstruction techniques, specifically Masked Autoencoders (MAE), to enable authorized users to recover features. In order to enhance activity recognition performance after feature recovery, MetAegisdevelops CTF-Net

(CNN-Transformer Fusion Network), a specialized activity classification network with multi-scale convolutional kernels designed to process the recovered features. The dual approach significantly improves activity recognition robustness for authorized users under the recovered feature, particularly in low signal-to-noise ratio scenarios.

We conducted extensive experiments under three physical world environments and multiple factors to verify the robustness and protection performance of MetAegis. The experimental results show that after deploying MetAegis, the accuracy of the sensing success rate of unauthorized users is less than 24% on average, while the average accuracy of authorized users remains above 88% under the regular deployment. Importantly, MetAegisverifies the impact on legitimate communication links. Experimental results show that under the UDP link throughput test, the impact on the throughput rate of the environmental communication link is less than 9%. In summary, our contributions include the points below:

- To the best of our knowledge, MetAegisis the first system to exploit the metasurface to simultaneously resist eavesdropping and protect authorized users' sensing ability on the channel side. In other words, the system only reshapes the existing electromagnetic wave environment and protects privacy-sensitive activities without any modification for the sensing devices.
- We proposed a novel obfuscation-recovery framework that can effectively obfuscate the activity features received by unauthorized users (regardless of location and number) while maintaining the accuracy and integrity of the legitimate user's sensing system.
- Finally, we implement the MetAegisend-to-end hardware prototype and validate its effectiveness in multiple real-world practical scenarios.

2 RELATED WORK

Wi-Fi Sensing Systems. With the broad deployment of Wi-Fi devices, it has become one of the most promising devices to assist wireless sensing. They claim a lot of exciting applications in Smart home, Smart health, and so on. For example, some prior works [13, 24, 67] aim to explore the feasibility of Wi-Fi based activity/gesture recognition. They built a strong relationship between the disturbed Wi-Fi signal and human activity by physical mathematics modeling and Deep Learning Network (DNN). Besides the activity recognition, some of them consider using Wi-Fi devices to monitor the daily vital signals like breathing rate [8], heart rate [15], and even the abdominal fat [42]. Furthermore, some interesting work is also involved in Wi-Fi sensing technology like indoor positioning [11, 32, 65] and humidity sensing [10, 17].

Wireless sensing protection technology. To protect against wireless sensing attacks, current research focuses on different objectives. As shown in Table 1, we analyze these approaches based on deployment methods, support for legitimate sensing, and the need for extra RF chains (i.e., the RF components that support antenna transmission). Specifically, Phycloak [44] and RF-Protect [49] attempt to achieve encryption by modifying the original signal. This involves using carefully designed hardware (such as full-duplex antennas) to tamper with the signal on the channel side to obfuscate illicit sensing links. However, this approach requires an additional RF chain, which undoubtedly increases protection costs and deployment burdens. Therefore, protection methods without additional RF chains are more effective. To this end, MIMOCrypt [35] and WiShield [18] utilize the transmitter's multiple antennas to scramble the long training sequence (LTS), preventing eavesdroppers from deciphering the CSI signal carrying sensing information. In addition, they support the legitimate link sensing. However, this approach relies on the number of transmitter antennas, which may limit the scenario of protection. To complement source-side approaches, the metasurface is considered an effective method for enhancing channel-side sensing security. IRShield [51] shows the feasibility of using the metasurface to obfuscate leaked Wi-Fi signals to prevent eavesdropping. Unfortunately, it obfuscates the whole wireless media which does not consider the case of authorized sensing users. In contrast, MetAegisis the first to utilize the metasurface to protect against eavesdroppers while supporting legitimate sensing. It's significant to note that since the metasurface's electronic

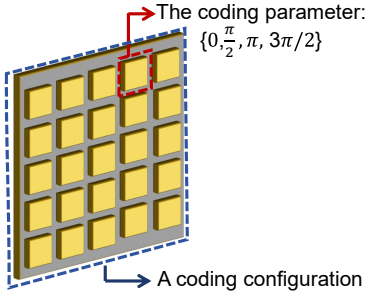


Fig. 2. The schematic of metasurface coding configurations.

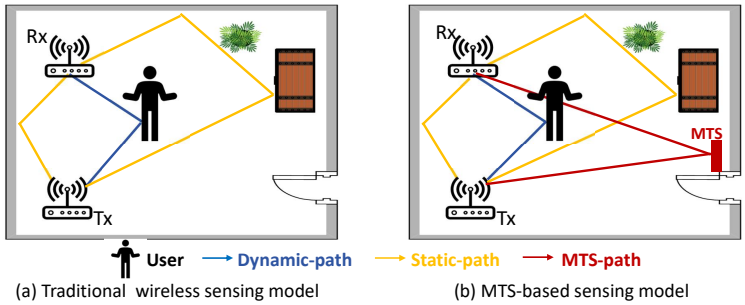


Fig. 3. The schematic diagram of the traditional and equipped metasurface spatial multipath channel propagation model.

components consist solely of PIN diodes, without any additional RF components, it can be easily deployed without the burden of extra RF chains. Furthermore, MetAegisis independent of protocols and source hardware requirements, eliminating the hardware requirements for the protected environment. Finally, MetAegisseamlessly integrates with existing source encryption solutions (such as MIMOCrypt) to achieve more effective wireless sensing protection.

Smart Metasurface Applications in Security. In recent years, metasurfaces have been proposed for some applications related to the security of wireless sensing and communication by reshaping the wireless channel. Specifically, Protego [27] considers injecting the random chaotic phase noise to confuse the eavesdroppers in side lobes. On the contrary, MetaFly [48] exploits a passive printed metasurface to achieve the adversarial communication link reshape. In addition, Hydra [66] can reduce the OFDM system's throughput by the carefully designed harmonic generation assisted by a metasurface. To the best of our knowledge, MetAegisis the first work to put the metasurface into wireless sensing security, which aims to prevent eavesdroppers and maintain the performance of authorized sensing users at the same time.

3 PRELIMINARY

In this section, we first introduce the principle of the programmable metasurface, followed by the mathematical model of the Wi-Fi-based wireless sensing method with metasurface.

3.1 Programmable Metasurface

Programmable metasurface has become a key technology for next-generation wireless networks, dynamically manipulating electromagnetic propagation by functioning as reconfigurable "radio prisms" through precisely engineered sub-wavelength unit-cells. These synthetic 2D arrays integrate low-cost tunable reflectors that collectively modify the phase and/or amplitude characteristics of incident waves to enable spatially coherent beam manipulation. A typical implementation consists of an $M \times N$ array of unit-cells, each with digital discrete phase-shifting capabilities. As illustrated in Figure 2, a 2-bit unit-cell provides four discrete phase compensation values (i.e., 0, $\pi/2$, π, and $3\pi/2$). Each of these phase states is referred to as a *coding parameter*, and the set of coding parameters for all the unit-cells is also called a *coding configuration*. Unlike active transceivers, the metasurface operates without power amplifiers or RF-chains by passively reflecting ambient signals, achieving low energy efficiency while maintaining low hardware complexity. This passive but programmable paradigm has catalyzed their adoption in 6G, with emerging applications extending beyond communications [28] to include RF sensing enhancement [30], indoor localization [23], and so on.

3.2 WiFi-based Sensing Model with Metasurface

During multipath propagation, Wi-Fi signals undergo dynamic variations due to changes in the sensing domain, such as human activities or respiration, resulting in changes in Wi-Fi channel parameters (i.e., amplitude, phase, frequency). The Channel State Information (CSI) describes the behavior of signal propagation from the transmitter to the receiver across multiple OFDM subcarriers. CSI can be extracted using commodity Wi-Fi Network Interface Cards (NICs), such as the Intel 5300 or Atheros 9580. Suppose $X(f,t)$ and $Y(f,t)$ represent the transmitted and received Wi-Fi signals on the sub-carrier of OFDM with frequency f and time t. Then, we have,

$$Y(f,t) = H(f,t) \times X(f,t), \tag{1}$$

where $H(f,t)$ is the CSI matrix which includes different multi-paths reflected to the receiver as shown in Figure 3(a), and it can be modeled according to [37]:

$$\begin{aligned} H(f,t) &= H_s(f,t) + H_d(f,t) \\ &= H_s(f,t) + \sum_{p=1}^{P} a_p(f,t)\, e^{-j2\pi f \tau_p(t)}, \end{aligned} \tag{2}$$

where $H_s(f,t)$ represents the static link components like the Light-of-Sight (LoS) between the transceiver, while $H_d(f,t)$ represents the dynamic path caused by the sensing object (human activity).

Recently, programmable metasurfaces have been introduced into the wireless sensing system due to their excellent capability to reconfigure the radio frequency (RF) environment [23, 30], as shown in Figure 3(b). The metasurface creates new "Tx-MTS-Rx" links in the sensing environment. Therefore, the sensing model in Eq. 2 can be rewritten as:

$$H(f,t) = H_s(f,t) + H_d(f,t) + H_m^{c_i}(f,t), \tag{3}$$

where $H_m^{c_i}(f,t) = \alpha_m^{c_i}(f,t)e^{-j\phi_m^{c_i}(f,t)}$, $\alpha_m^{c_i}(f,t)$ and $\phi_m^{c_i}(f,t)$ are the amplitude and phase of the incident signal reconfigured by the i^{th} coding configuration on metasurface and finally propagated to the receiver.

4 OVERVIEW

4.1 Threat Model

Threat Scenarios: Consider a wireless sensing system comprising a transmitter-receiver pair (Alice and Bob), which could represent practical devices such as Wi-Fi access points, mobile terminals, or laptops. An adversary (Eve), monitors the wireless environment and can infer sensitive information through advanced analysis of CSI variations, including but not limited to human activities (e.g., movement patterns, gait characteristics), and physiological information (e.g., respiration). Critically, this security threat operates covertly, and legitimate users are typically unaware of such surveillance. For example, in elderly care centers equipped with smart WiFi-based fall detection systems such as CareClaim [2], potential illegal attackers could monitor and identify vulnerable states, such as falling, to facilitate targeted crimes, including burglary [5].

System Goal: MetAegisis designed to protect intelligent Internet of Things (IoT) environments, particularly in privacy-sensitive scenarios such as smart homes and elder care monitoring. Neither the number nor the locations of the adversaries (i.e., Eves) are known to MetAegis. MetAegisis designed with two core objectives:

(1) **Obfuscation for Wireless Sensing Links.** By disrupting CSI between Alice and unknown Eves across the entire spatial domain, MetAegiseffectively thwarts eavesdropping attempts by obfuscating the received feature patterns, minimizing the risk of sensitive information leakage.

(2) **Sensing for Authorized Users.** MetAegisensures accurate recovery of activity-related features between the authorized sensing links (Alice and Bob), allowing legitimate users can maintain high sensing accuracy by leveraging physical-layer signal characteristics.

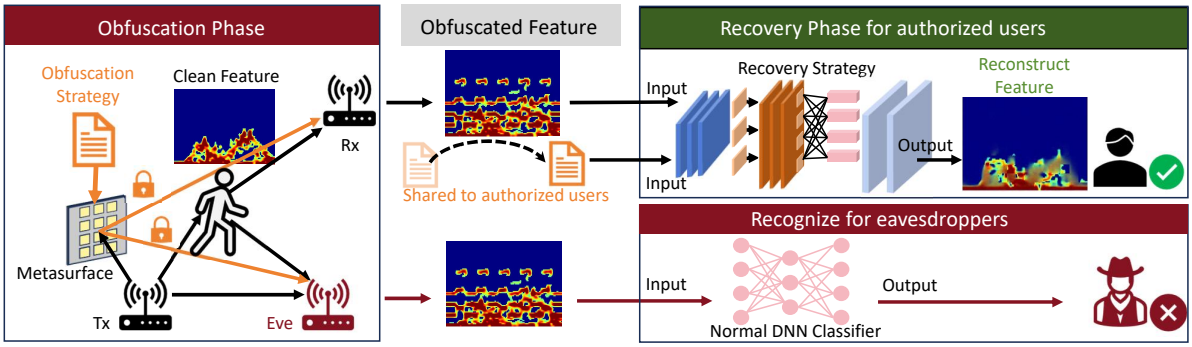


Fig. 4. System overview of MetAegis.

4.2 System Overview

The system workflow is shown in Figure 4. In order to defend against eavesdroppers, we analyze the key insight of disturbance that the association between duty cycle rate and harmonic components frequency in Section 5.1.1. Based on it, we proposed the obfuscation strategy in Section 5.1.2. Then, we proposed a coding configuration set strategy to expand the defense to multiple directions in Section 5.2. To maintain the performance of authorized sensing, we recover the activity-related features and propose an activity recognition enhancement scheme to ensure legitimate users' performance in Section 5.3. Ultimately, under MetAegis's protection, eavesdroppers will be confused, and authorized users can maintain normal sensing tasks.

5 METAEGIS: SYSTEM DESIGN

5.1 Single-angle Obfuscation for Fixed Adversarial Scenarios

First of all, we consider a fundamental use case where feature obfuscation is implemented along a single predetermined direction. That simplified scenario serves as the foundation for our subsequent analysis of the more complex obfuscation strategy against multi-directional cases.

5.1.1 Core Obfuscation Strategy Analyze. Tracing back to Eq. 3, it can be found that $H_m^{C_i}(f,t)$ will impact the whole $H(f,t)$, which is used to extract the activity feature. Therefore, a key intuition is whether the metasurface is exploited to confuse the feature. We use the Doppler frequency shift (DFS) spectrogram feature as an example: the feature from $H(f,t)$ can be calculated by short-time Fourier transform (STFT) [57]:

$$STFT(H(f,t)) = f_s(t) + f_d(t) + f_m(t) \tag{4}$$

where $f_s(t)$ is the frequency from the static path, typically considered as 0 due to without any changes. $f_d(t)$ represents the frequency shift caused by human motions (activities or gestures). $f_m(t)$ generated by the metasurface reflected path. It is worth noting that $f_m(t)$ will equal 0 as same as $f_s(t)$ if the coding configuration is stable (which can be regarded as an extra static link from transmitter to metasurface to receiver).

Therefore, when $f_m(t) \neq 0$, the frequency components introduced by the metasurface link can significantly alter activity-related features in the frequency domain. A straightforward obfuscation idea, as shown in Figure 5, is to generate a stable interference frequency component by switching coding configurations at a constant rate. This approach leverages the amplitude variations in the CSI signal caused by different coding configurations [22] to confuse the activity feature. However, human activity patterns typically exhibit multiple frequency components within any given time window. Consequently, a single adversarial frequency component cannot fully disrupt the feature spectrum, and the eavesdropper is able to detect or remove it easily.

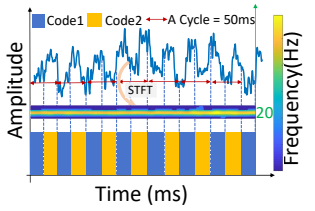


Fig. 5. CSI results under switching codings.

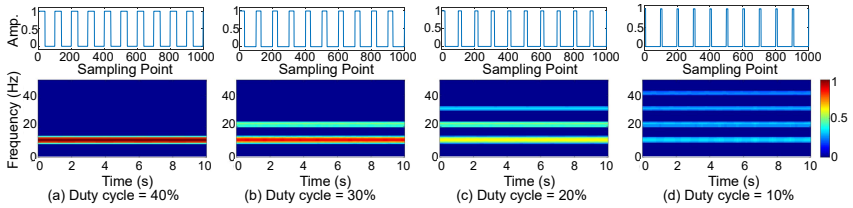


Fig. 6. Frequency domain results at different duty cycles.

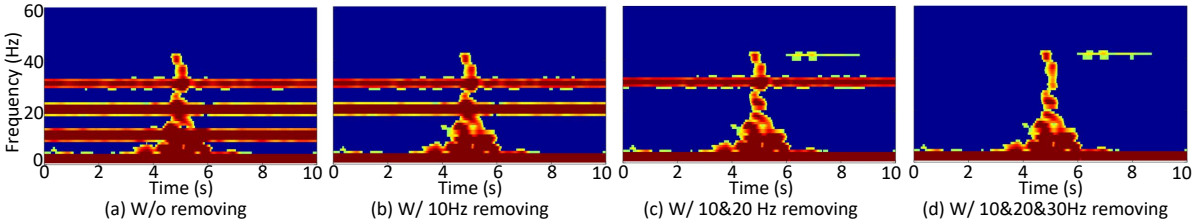


Fig. 7. The results of in different obfusration removing.

Fortunately, inspired by prior work that demonstrates how different duty cycles affect harmonic interference in common-mode conducted emissions [40], we proposed a hypothesis that we can exploit the harmonics to achieve multiple frequency confusion by operating the duty cycle rate. To verify our hypothesis, we can formulate the signal $S(t)$ in the time domain as a square wave, and the expression is as follows:

$$S(t) = \begin{cases} A_1 & \text{if } 0 \le t < T_{high} \\ A_2 & \text{if } T_{high} \le t < T \end{cases} \tag{5}$$

where A_1 and A_2 are the signal amplitudes of the metasurface generated by different coding configurations. It can be calculated by: $A_i = \left|\alpha_{code_i} e^{-j\phi_{code_i}}\right|, i \in \{1, 2\}$. T_{high} and T represent the high amplitude duration time and the whole cycle of $f(t)$, respectively. Exploiting the Fourier Series, $S(t)$ can be rewritten as:

$$S(t) = a_0 + \sum_{n=1}^{\infty} [a_n cos(\frac{2\pi nt}{T}) + b_n sin(\frac{2\pi nt}{T})] \tag{6}$$

where a_0 is the DC component, which can be ignored. a_n and b_n represent the Cosine and Sine Term Coefficients. They can be calculated by $a_n = \frac{2}{T}\int_0^T f(t)cos(\frac{2\pi nt}{T})d_t$ and $b_n = \frac{2}{T}\int_0^T f(t)sin(\frac{2\pi nt}{T})d_t$. After simplification, the amplitude coefficient of the harmonic C_n can be represented as:

$$C_n = \frac{2|A_1 - A_2|}{n\pi} |sin(n\pi \frac{T_{high}}{T})| \tag{7}$$

where $\frac{T_{high}}{T}$ is the duty cycle rate D. Based on the above analysis, MetAegiscan generate a multi-frequency coverage to confuse the original feature in the frequency domain by adjusting the duty cycle rate to control the harmonics. We use a case validation experiment employed duty cycle rates ranging from 40% to 10% in steps of 10%, with results depicted in Figure 6. We have the following key observations: When we switched the coding configurations at different delays, the high-frequency components appeared differentiated. In addition, as the duty cycle rate decreased, ideally all frequency components exhibited amplitude attenuation as analyzed in Eq. 7. Therefore, MetAegiscan achieve multiple frequency confusion exploiting the suitable duty cycle rate controlling.

5.1.2 Obfuscation Strategy. While the previous subsection demonstrated the feasibility of multi-frequency perturbations injection, due to the time-regular nature (perturbations at a single frequency last for the entire

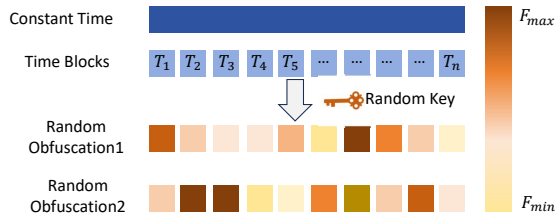


Fig. 8. Diagram of the dynamic random obfuscation.

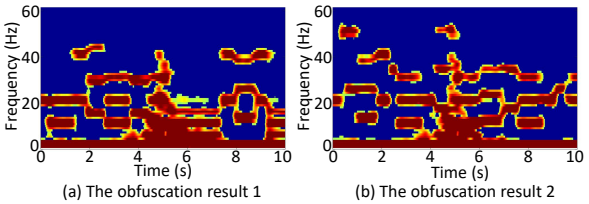


Fig. 9. The simulation results of random obfuscation.

time) of the basic obfuscation strategy, the eavesdroppers can remove the perturbations by observing the received signal several times. To verify the potential concerns, we adjust the duty cycle to add the perturbations to the activity feature as shown in Figure 7(a). Then, we simulate an eavesdropper deploying 10 Hz single-band filter, 10&20 Hz dual-band filters, and 10&20&30 Hz multi-band filters to filter the signal. The results are shown in Figure 7(b-d). It demonstrates that such static multi-frequency perturbations can be effectively removed through simple filtering operations, enabling eavesdroppers to recover clean activity features.

To enhance the obfuscation strategy, we propose a dynamic random obfuscation strategy. As illustrated in Figure 8, MetAegis first splits the time domain into multiple sub-blocks. Then, a Random Key is utilized to generate different frequency perturbations within each sub-block.

Random Key Design: For the discrete time blocks $\{T_1, T_2, ..., T_n\}$, the Random Key Ψ firstly determins the maximum frequency F_{max} and minimum frequency F_{min}. Then the Random Key Ψ can be defined as:

$$\Psi(i) = \begin{cases} rand(F_1, P_i) & F_{min} \le F_1 < F_{bound}, P_i \in P_{bound} \\ rand(F_2, P_j) & F_{min} \le F_2 < F_{max}, P_j \in C_u^{P_{bound}} \end{cases} \tag{8}$$

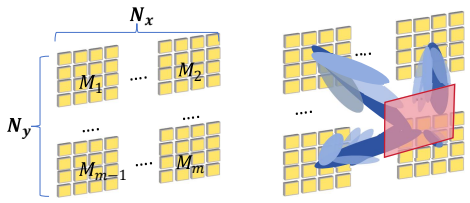
where F_{bound} is the max frequency of daily activity features (about 40 Hz) and P_{bound} is the probability of falling into the valid area (we define it as 0.7). It means that the generated random frequency perturbations can fall into $[F_{min}, F_{bound}]$ to ensure the effective perturbations. As the simulation results show in Figure 9, the different obfuscations can be generated as the Random Key Ψ changes. Crucially, since the eavesdroppers have less prior knowledge of Ψ, the perturbations become fundamentally irremovable through conventional filtering techniques.

5.2 Multi-angle Obfuscation for Dynamic Adversarial Scenarios

Building upon our successful demonstration of single-angle feature obfuscation, now we consider extending the framework to achieve comprehensive multi-directional feature obfuscation. This consideration aims at defending the more complex and practical threat scenario where exists multiple potential adversaries located at multi-positions around the sensing space.

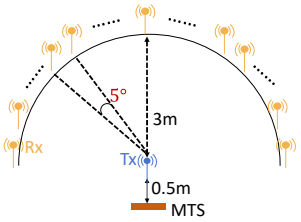
Backtracking to the Eq. 5, the beamforming coding configurations $BFCoding_{\theta_1}$ can be used to provide the high amplitude A_1 in direction θ_1. Then other coding configurations $BFCoding_{\theta_2}$ can provide the low amplitude A_2. It's easy to find that both θ_1 and θ_2 will be obfuscated due to the reciprocity of the high amplitude and low amplitude caused by the beamforming coding configurations. Thus, it seems feasible to achieve multi-angle obfuscation by scanning all beamforming configurations across the metasurface's field of view (FoV, typically in the range of $[-60°, 60°]$ [28]) in principle. Unfortunately, each beam width is only about 10° in practice. It means that we need 13 coding configurations to cover all the directions at least! From the results shown in Figure 6, the amplitude of different harmonics will decrease as the duty cycle rate decreases. It implies that the obfuscation signal fails to effectively mask the original activity features in the frequency domain, enabling eavesdroppers to recover clean signatures through spectral analysis.

Based on the above analysis, the selected duty cycle rate of switching coding configurations significantly impacts perturbation performance. Besides, the beams provided by coding configuration sets must cover all

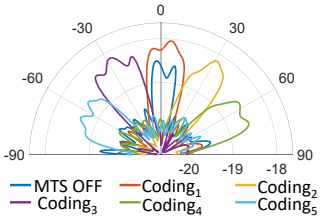


(a) The idea of sub-arrays (b) Schematic of Beam-broaden

Fig. 10. The idea of the Beam-broaden.



(a) The layout of experiment



(b) The results of different codings

Fig. 11. The results of the Beam-broaden.

directions within the FoV of the metasurface. Firstly, we set the goal duty cycle rate at 20% to achieve the most effective obfuscation. Furthermore, to achieve a trade-off between duty cycle rate and directional coverage, it is necessary to use only 5 coding configurations to cover multiple directions (each coding configuration covers about 30°). To achieve this goal, inspired by the **Beam-broaden** technology in MIMO antenna array [38, 45], we introduce a sub-array idea for the metasurface.

Beam-broaden scanning. Specifically, as shown in Figure 10(a). We try to divide the metasurface with $N_x \times N_y$ unit-cells into M sub-arrays which equipped $N_{\frac{x}{\sqrt{M}}} \times N_{\frac{y}{\sqrt{M}}}$ unit-cells. We aim to configure each sub-array individually to achieve nearby beam steering for the desired direction θ_0, as shown in the diagram in Figure 10(b).

To achieve this goal, firstly we need to determine the beam directions for each sub-array. It can be formulated as $\{AF_{\theta_{M_1},\varphi_{M_1}}, AF_{\theta_{M_2},\varphi_{M_2}} ....AF_{\theta_{M_m},\varphi_{M_m}}\}$. Where $AF(.)$ is the array factor of which can be expressed by:

$$AF = \sum_{m=1}^{\sqrt{M}} \sum_{n=1}^{\sqrt{M}} a_{(m,n)} e^{j[\phi^T_{(m,n)} - \phi^I_{(m,n)}]_{2-bit}} \tag{9}$$

$$\phi^T_{(m,n)} = -k\,(x_m \sin\theta_l \cos\varphi_l + y_n \sin\theta_l \sin\varphi_l) \tag{10}$$

Where $\phi^T_{(m,n)}$ is the theoretical phase distribution. x_m and y_n is the distances of the $(m,n)^{th}$ unit-cell of sub-arrays relative to the origin of coordinate, $a_{(m,n))}$ is the amplitude. $\phi^I_{(m,n)}$ is the arrival-phase of electromagnetic wave impinges the $(m,n)^{th}$ unit-cells. Due to the discrete phase compensation, the phase needs to change into discrete phase in 2-bit (the metasurface prototype we used is 2-bit).

To search effective $(\theta_{M_i}, \varphi_{M_i})$, we use particle swarm optimization (PSO) to optimize the configuration. We define the metasurface divided into 4 sub-arrays and the goal of optimization in each iteration as:

$$\phi^{T_{Broaden}}_{(N_x,N_y)} = \begin{bmatrix} \phi^T_{(M_1,M_1)}|_{(\theta_{M_1},\varphi_{M_1})} & \phi^T_{(M_2,M_2)}|_{(\theta_{M_2},\varphi_{M_2})} \\ \phi^T_{(M_3,M_3)}|_{(\theta_{M_3},\varphi_{M_3})} & \phi^T_{(M_4,M_4)}|_{(\theta_{M_4},\varphi_{M_4})} \end{bmatrix} \tag{11}$$

in order to promise the center frequency (θ_0, φ_0), each sub beam $(\theta_{M_i}, \varphi_{M_i})$ is restricted into $\{\theta_0, \varphi_0\} - \frac{BW}{2} \leq \{\theta_{M_i}, \varphi_{M_i}\} \leq \{\theta_0, \varphi_0\} + \frac{BW}{2}$. Where the BW is the desired beam width for beam-broadening, and we set it equal to 30°. We aim at broadening the beam without any nulling in the desired direction, thus, the whole objective function of optimization is as follows:

$$\Im \in min(Var(\sum_{m=1}^{N_x} \sum_{n=1}^{N_y} a_{(m,n)} e^{j[\phi^{T_{Broaden}}_{(m,n)} - \phi^I_{(m,n)}]_{2-bit}})) \tag{12}$$

We conducted an experiment to demonstrate the effectiveness of the Beam-broaden design. To promise the duty cycle, we generate 5 coding configurations, which include the −60°, −30°, 0°, 30° and 60°, and the deployment is shown in Figure 11(a). The approach effectively broadens the beam width as shown in Figure 11(b), allowing MetAegisto achieve full-FoV multi-directional obfuscation with just five beam coding configurations.

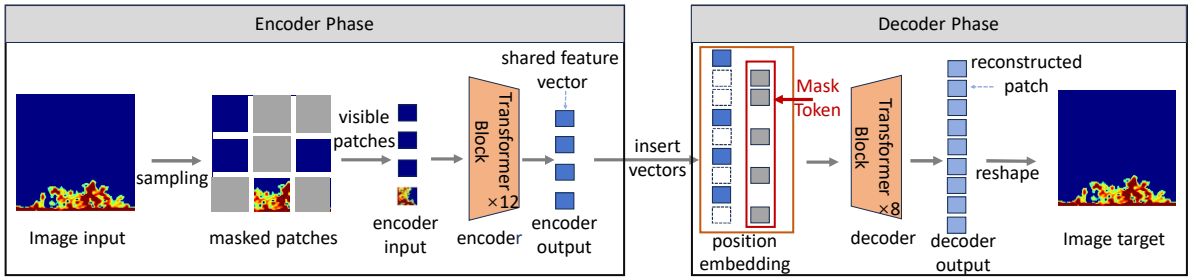


Fig. 12. The pipeline of MAE.

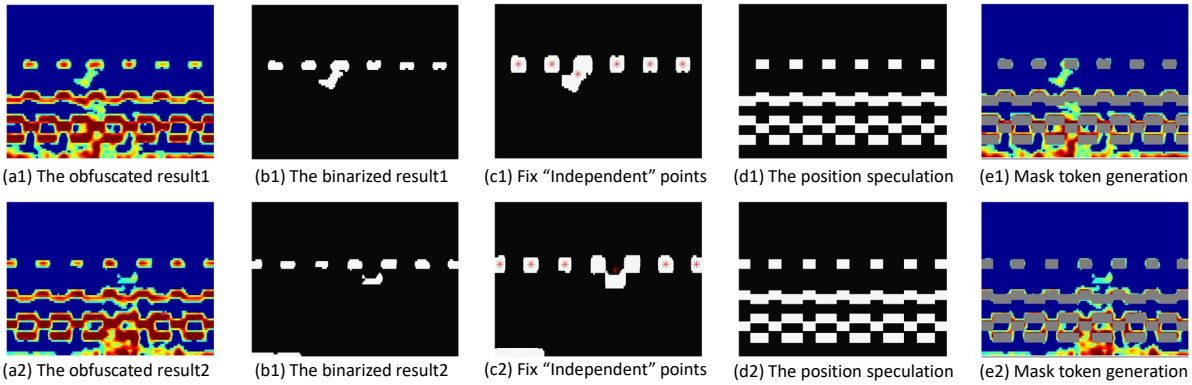


Fig. 13. The processing of the mask token generation from the encrypted feature received by authorized users.

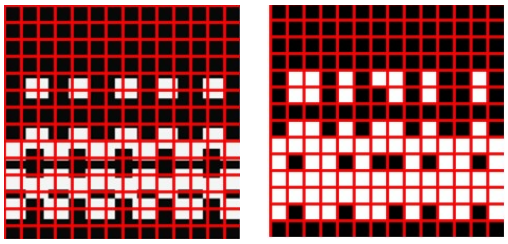
5.3 Sensing Recovery Framework for Authenticated Users

To enable authorized users to maintain sensing ability under the obfuscated feature caused by MetAegis, we focus on recovering the activity-related features from the obfuscated signals in the frequency domain.

5.3.1 Framework Overview. As mentioned in Section 5.1, the key method of obfuscation concentrates on covering the original feature by multiple harmonic components. The disturbance added to the feature like a "mask". It gives us an intuition that we can transfer the recovery task to a masked image restoration task in the Computer Vision (CV) field. Therefore, we borrowed the advanced visual restoration framework Mask Autoencoders (MAE) [16] to recover the activity-related features in the frequency domain. It is worth noting that we concentrate more on the feasibility of recovering the obfuscated feature, and the method is not exclusive, other state-of-the-art reconstruction techniques can also be adapted to incorporate our core idea.

The pipeline is shown in Figure 12. Specifically, we leverage the ViT-Base architecture [12] to recover the activity-based features, which is pre-trained on the ImageNet-1K dataset, acquiring strong general feature representation capabilities and providing an excellent parameter initialization for downstream tasks. Building upon the pre-trained ViT-Base model, we perform domain-specific adaptation using time-frequency activity-related features (850 samples per activity, total 4250 data). Then, the model can extract low-dimensional features from obfuscated time-frequency inputs using the encoder and recover complete feature maps through the decoder and the mask tokens.

5.3.2 The Generation of Mask Tokens. Accurate positions of mask tokens, as shown in Figure 13(a), are crucial to recovering the activity-related activities while using the decoder. Although authorized users can access the mask sequence used during obfuscation, due to each feature extraction interval not strictly synchronized with the



(a) The result of mask division (b) The adapted mask token

Fig. 14. The results of the original mask division and the mask token after the adaptation strategy.

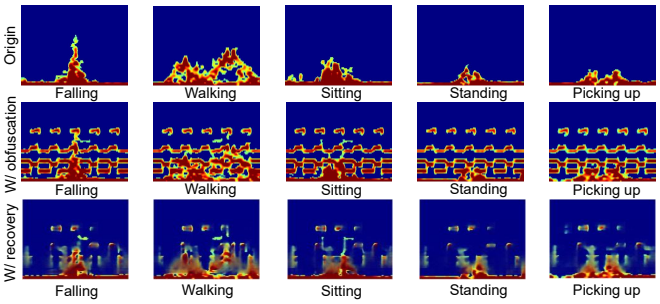


Fig. 15. The results of different activities' feature without obfuscation, with obfuscation and with recovery.

switching of the metasurface configuration (strictly synchronized may incur the cost and limitations deployed MetAegis), which introduces positional uncertainty.

To address the problem, our key insight is to leverage the independence of high-frequency obfuscation masks, which exhibit clearer spatial separation from low-frequency activity-related features. Specifically, here are two cases of features covered by asynchronous masks as shown in Figure 13. First of all, we separate the high-frequency masks (we defined them as "independent points") as shown in Figure 13 (b1,b2). Besides, we apply morphological dilation to regularize those shapes of independent points by filling gaps and connecting adjacent regions, as shown in Figure 13 (c1,c2). Once we fixed these independent points, the legitimate user can speculate the location of the whole mask, leverage the prefabricated mask sequence generated by the Random Key (known only to transmitters and authorized users). The results of the mask location speculation and mask generation under asynchronous conditions are shown in Figure 13 (d1,d2) and Figure 13 (e1,e2). They show that, based on the method we proposed, the mask can be located reliably under asynchronous conditions to support MAE framework recovery.

To effectively utilize pre-trained models while maintaining their learned universal feature representations, we maintain the standard 16×16 input specification consistent with ViT-Base's original configuration. However, when we adjust the mask to the same size, as shown in Figure 14(a). It can be found that the mask position cannot be fully matched with the grid of the specification parameters, which implies that we cannot add a mask to guide the model to learn the masked position. To address the limitation, we propose an adaptive adjustment strategy to reconstruct the mask token. Specifically, The whole mask $M_{gloabal}$ after size adjustment can be refered as $M_{gloabal} \in \{0,1\}^{224\times224}$. We divided the $M_{gloabal}$ as $(P \times P)^{th}$ sub-areas, where P represents the size of specification parameters (here equals 14). Therefore we calculate the mask rate for each sub-areas $B_{i,j}$ as:

$$\rho_{i,j} = \frac{1}{P^2} \sum_{x=1}^{P} \sum_{y=1}^{P} B_{i,j}(x,y) \tag{13}$$

The larger the value of $\rho_{i,j}$ is, it emplies the more the mask fills the sub-area. In this way, we adaptively adjust the mask tokens by discretizing the mask of each sub-area with respect to $\rho_{i,j} > T_r$, where we choose the threshold $T_r = 0.25$. Finally, the adapted mask token M^{token} can be calculate by:

$$M_{i,j}^{token} = \begin{cases} 1, & if \rho_{i,j} > 0.25 \\ 0, & otherwise \end{cases} \tag{14}$$

The result of the M^{token} can be shown in Figure 14(b), which satisfies the input requirement of the size limitation of specification parameters.

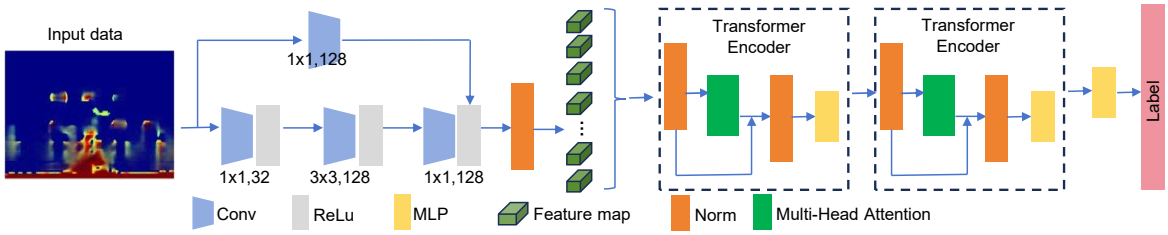


Fig. 16. The structure of proposed framework.

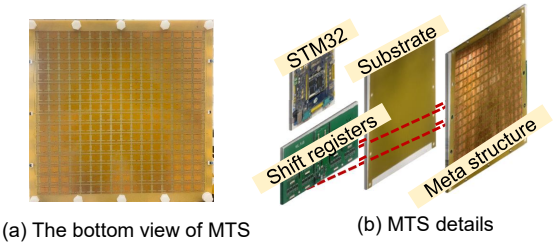
5.3.3 Recognition Enhancement Design. The recovered activity-based features using the MAE scheme are shown in Figure 15. We can see that the time-frequency activity-based features have been recovered, but it still has some remaining mask noise. To enhance recognition accuracy, we propose a lightweight CNN-Transformer hybrid network, as shown in Figure 16. The multi-scale CNN module employs varying kernel sizes to capture both local details (small kernels) and extended noise patterns (large kernels), while the Transformer module uses self-attention to adaptively emphasize important features while suppressing noisy regions, improving global modeling capability, and enhancing noise robustness.

The overall framework takes the time-frequency feature image recovered from the MAE decoder as input, with dimensions $(224, 224, 3)$. A dual-path CNN module efficiently extracts local and global features using a combination of multiple 1×1 and 3×3 convolutional layers, producing a discriminative $(7, 28, 128)$ feature representation, where 128 represents the number of feature channels, and the spatial resolution is 7×28. The multi-scale feature representation $(7, 28, 128)$, is rearranged $(196, 128)$ through spatial flattening, reshaping, and permutation to match the Transformer input format. Then Transformer Encoder, consisting of 2 layers of Transformer Encoder Layer (with model dimension d_model= 128, 8 attention heads, and a feedforward dimension of 512), processes that rearrange the feature matrix. Each layer applies multi-head self-attention, layer normalization, and a feedforward network.

The Transformer-encoded feature representations are fed into a multi-layer perceptron (MLP) classifier for final activity recognition. The classifier realizes the projection mapping from high-dimensional feature space to category labels through a fully connected layer, uses cross-entropy loss, and iteratively optimizes the network parameters through the gradient descent algorithm to enable the model to accurately capture the semantic activity-based features. In order to further enhance the model's adaptability to noise after MAE recovery, this paper leverages a noise augmentation strategy during the training phase. Specifically, artificial noise is added to the time-frequency activity-related feature map, and mainly includes random scattered noise, block noise and edge blur three types. The strategy enables the model to learn more prior knowledge about noise during training, improving the robustness of feature extraction. Furthermore, we augment the training set with 10% MAE-recovery features into the pre-collected clean dataset, which can release the imbalance between the reconstructed feature and clean feature, reducing the training cost for users without recollecting the dataset.

6 IMPLEMENT

Metasurface prototype. We construct a prototype of MetAegis's metasurface that consists of 16×16 array of unit-cells. These unit-cells are uniformly distributed inside an area of $32 \times 32cm^2$ and a thickness of $6.2mm$, as shown in Figure 17(a). To control the PIN diode (SMP1340-040LF [7]) states of each unit-cell, the power consumption of each PIN diode is only $8.8mW$ (0.01A*0.88V). A bias layer is integrated to deliver DC bias voltage (0 V/5 V) to the PIN diodes. Therefore, each unit-cell functions as a 2-bit phase shifter which provides four discrete phase states (i.e., 0, $\pi/2$, π, and $3\pi/2$). We carefully optimize both the material composition and geometric design of the unit-cells to ensure operation in both the 2.4 GHz and 5 GHz ISM bands. Meanwhile, the substrate costs approximately 160$ and each PIN diode costs 0.17$, and the total material cost is about 287$.



(a) The bottom view of MTS (b) MTS details

Fig. 17. The metasurface prototype.

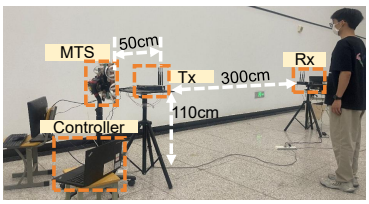


Fig. 18. The experiment setup.

Hardware control. In order to configure the metasurface, as shown in Figure 17(b), we design a control circuit module consisting of a microcontroller (i.e., STM32H743IIT6, STMicroelectronics) and 64 SN74LV595 shift registers to provide different DC voltages (0 V/5 V) for each unit-cell. Specifically, we divide the entire metasurface board into 4 zones. For each zone, we use two channels in the microcontroller to transmit a data stream with 128 bits to control 128 PIN diodes. Via the above setup, the controller is now able to independently configure the state of each unit-cell's PIN diode.

7 EVALUATION

7.1 Experimental Setup

In this section, we introduce the experimental setup and the evaluation metric.

Wireless setup. We achieve the Wi-Fi sensing transceiver prototype equipped with Intel 5300 NIC and Qualcomm-Atheros AR9580 NIC. The transmitter is equipped with 1 Omnidirectional antenna, and the receiver is equipped with 3 Omnidirectional antennas. We install the Linux CSI Tools [3] on the receivers to collect CSI measurements for evaluation. The transmission rate of the system is 1000 packets per second. As shown in Figure 18, all transceivers are placed at a height of 110 cm so that the motion of users with different heights can be detected. For the details of evaluations, we keep the distance between the transceiver at 3 m as the default setting, and the metasurface is deployed behind the transmitter at 0.5 m.

Participants. We totally recruit 7 volunteers (5 male and 2 female), with heights ranging from 161 cm to 189 cm and weights ranging from 48 kg to 92 kg to repeat 5 daily activities (including sitting, walking, falling, picking up, and standing).

Data preprocessing. We borrowed the signal processing method in [41].Specifically, we first exploit the Discrete Wavelet Transform (DWT) to remove the high-frequency noise (which doesn't include the activity information) in the CSI signal. Then, we choose principal component analysis (PCA) to decorrelate the CSI streams. Finally, we apply STFT on the selected principal components individually and average the spectrograms to obtain a single spectrogram for each activity.

Evaluation metric. In order to completely evaluate the MetAegisperformance for preventing against eavesdroppers and recovering authorized users. We choose **Accuracy** to evaluate in a series of experiments. The Accuracy is calculated by: $Accuracy = \frac{N^{cor}}{N^{all}}$, with N^{cor} and N^{all} denoting the number of correctly recognized and the number of all activities.

Naming conventions. We exploit three representations to distinguish the eavesdroppers and legitimate users: **(1) w/o MetAegis:** which represents the free baseline, and the situation without MetAegisdeployment. It implies that both authorized users and eavesdroppers can receive the same clean feature without any protection mechanism. **(2) Authorized users:** It represents the full MetAegispipeline, and the authorized user exploits the method proposed by MetAegisto recover the original feature when receiving the obfuscated feature, which concentrates on evaluating the performance from the authorized user's view. **(3) Eavesdroppers:** which represents

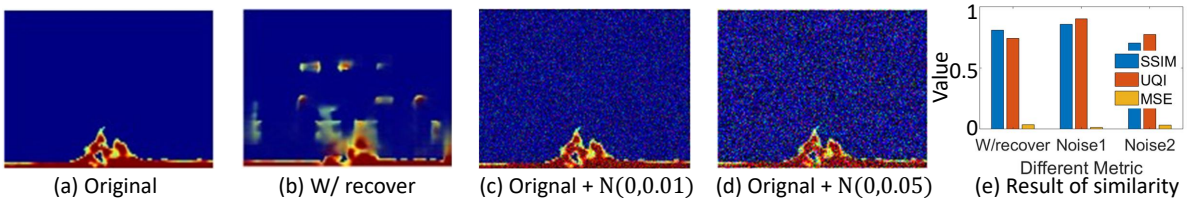


Fig. 19. The results of indicators evaluated the feature with recovering and two different benchmarks.

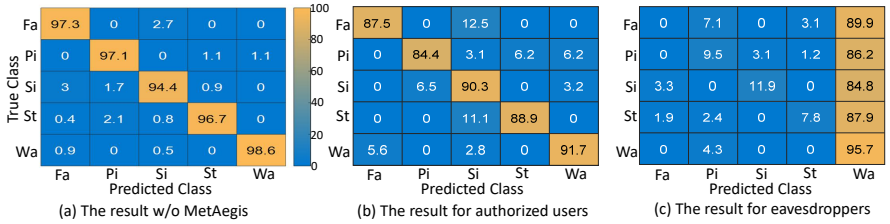
obfuscated but unrecovered, it imlies that the MetAegisobfuscation is applied but recovery is unavailable. It aims to evaluate the performance from the eavesdropper's view.

7.2 Micro-benchmark

The verification of feature recovery quality. To verify the recovery quality of the MAE framework for encrypted features, we use the three most popular indicators [47, 50] for evaluation, including Structural Similarity (SSIM), Universal Quality Index (UQI), and mean squared error (MSE). Specifically, as shown in Figure 19 (c-d), we use Gaussian noise with mean 0, variance 0.001 and 0.005 added to the clean feature as baselines to compare the value of indicators. The result is shown in Figure 19 (e), the values of three indicators are 0.807, 0.036, 0.739 for the recovery feature. It can be found that the SSIM value is between the two baselines, which means that the quality of the restored image is better than $Baseline^2$ and slightly lower than $Baseline^1$. It should be noted that the quality of the restored image is slightly lower than $Baseline^2$ in UQI and MSE. This is because SSIM is designed by modeling any image distortion as a combination of three factors that are loss of correlation, luminance distortion, and contrast distortion. However, UQI and MSE cannot correlate all subjective evaluations, resulting in instability [46]. But the totally quality is close to that of $Baseline^2$ from the three indicator, it can still reflect that the restored features can ensure high quality.

The verification of different activities under obfuscation and recovery methods. Five daily activities, including falling (Fa), standing (St), picking up (Pi), sitting (Si) and walking (Wa) are selected to evaluate the performance of MetAegisproposed obfuscation and recovery methods. We first establish a baseline by evaluating the performance using clean, unobfuscated features (which can simulate the eavesdroppers without any protection). Then we evaluate the performance when MetAegisstarts work. The experimental results reveal three main findings: (1) Vulnerability of unprotected sensing systems. As shown in Figure 20(a), the eavesdroppers can easily achieve eavesdropping. (2) Authorized User Performance. When deploying the metasurface, the authorized user can recover the feature and achieve above 88% accuracy on average, as shown in Figure 20(b), confirming effective feature recovery. (3) Eavesdroppers Resistance. In addition, the eavesdroppers can hardly recognize the activity due to without recovery method as shown in Figure 20(c). It is worth noting that almost all activities are recognized as walking. The reason is that the mask we apply will obfuscate the feature, the Deep Learning Model is harsh to match any feature from the activities except walking (due to its long duration). These results demonstrate that MetAegissystem can successfully achieves secure activity feature obfuscation while maintaining reliable sensing performance for authorized users.

Compared with other sensing encryption work based on the metasurface. To compare with other defending methods assisted by the metasurface. We compare MetAegiswith IRShield [51]. Specifically, we switch the metasurface configuration (following the strategy IRShield proposed) rapidly at 20 configurations/sec to randomize the channel. The result is shown in Figure 21. Both the MetAegisand IRShield demonstrated excellent performance in defending eavesdroppers, the average attack accuracy falling below 25%. However, benefiting from the fact that MetAegistakes legitimate sensing links into account, the legitimate user can maintain an average accuracy of 87.5%. In contrast, IRShield uses an "undiscriminating" approach, which does not distinguish



(a) The result w/o MetAegis (b) The result for authorized users (c) The result for eavesdroppers

Fig. 20. The results for obfuscation and recovery.

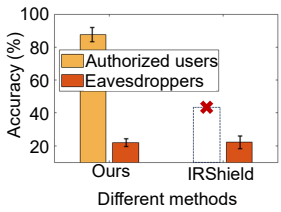
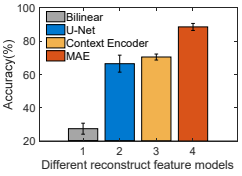


Fig. 21. Compare with related work.

Compared Model	Reconstruct insights
Bilinear	Local neighborhood linear weighting
U-Net	Encoder-Decoder architecture
Context Encoder	Combining reconstruction loss and adversarial loss
MAE (Ours)	Pre-training based on ViT and reconstruction fine-tuning

(a) The details of different compared models

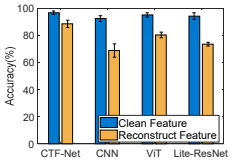


(b) The result of different models

Fig. 22. Verification of the recovery recovery model.

Compared Model	Architecture
CTF-Net (Ours)	1 Con-Layer + 2 Transformer + MLP
CNN	3 Con-Layer + 1 Max Pooling + FC-Layer
ViT	2 Transformer
Lite-ResNet	Residual Network + 1 Max Pooling

(a) The details of different compared models



(b) The result of different models

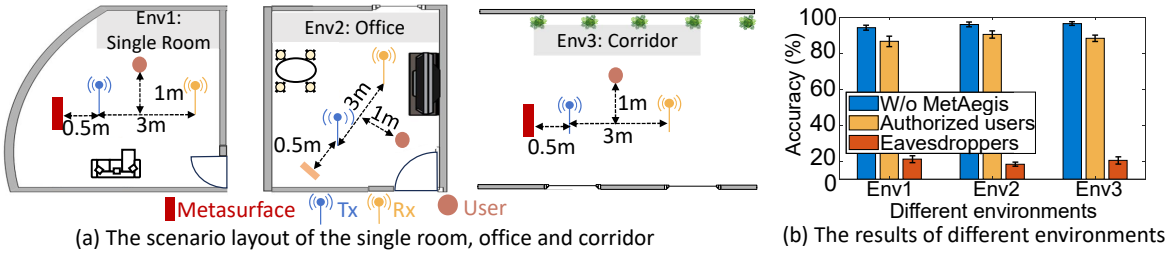
Fig. 23. Verification of the classification stategy.

between legitimate users and eavesdroppers; legitimate users cannot correctly utilize sensing links. Therefore, MetAegishas a wider range of applications.

Compared the reconstructed feature performance under the recovery phase of MetAegiswith other baselines. In order to validate the effectiveness of the feature recovery method for authorized users proposed in Section 5.3, specifically, we set up three basic baselines for comparison as shown in Figure 22(a): (1) Bilinear, which fills the missing area by using local neighborhood linear weighting; (2) U-Net, which uses an encoder-decoder architecture to achieve supervised image restoration; (3) Context Encoder, which combines self-supervision and adversarial training reconstruction capabilities; all methods are trained and tested on the same dataset. As demonstrated in Figure 22(b). MAE-based approach exhibits superior performance in the obfuscated time-frequency feature recovery task, achieving 88.6% classification accuracy, far exceeding the comparison method. The traditional bilinear interpolation method only achieves an accuracy of 27.4%; U-Net and Context Encoder achieve accuracies of 68.1% and 70.5% respectively. Although deep learning architecture outperforms traditional interpolation, it is limited by the local receptive field of the convolutional neural network and fails to effectively capture long-range context dependencies. In contrast, MAE can still restore the semantics of time-frequency domain features in high-mask scenarios, through the global attention mechanism and self-supervised pre-training strategy of the ViT architecture, verifying its ability to efficiently reconstruct obfuscted areas.

Compared the recognition model performance of MetAegisand other baselines. In this experiment, we aim to verify the effectiveness of the classification strategy of MetAegisby conducting comparative experiments with the following three baselines: 1) CNN, 2) A Vision Transformer (ViT) model, and 3) A Lite-ResNet model. The architectural details of these models are illustrated in Figure 23(a). During the evaluation phase, the same dataset is employed for all models and their performance under two conditions: processing clean features and handling the reconstructed features after the recovery strategy. As demonstrated in Figure 23(b), all models achieve classification accuracy above 90% when processing clean features. However, significant performance variations emerge when dealing with reconstructed features, with the average accuracy of 88.6%, 68.9%, 80.4% and 73.6%, respectively. Notably, The CTF-Net shows the more superior performance due to its enhanced robustness when processing noisy image and its ability to effectively capture multi-scale features. The experimental results demonstrate that the classification strategy of MetAegissystem can reliably maintain high performance for authorized users during feature recovery, demonstrating the feasibly in practical scenarios.

.



(a) The scenario layout of the single room, office and corridor

(b) The results of different environments

Fig. 24. The performance of different environments.

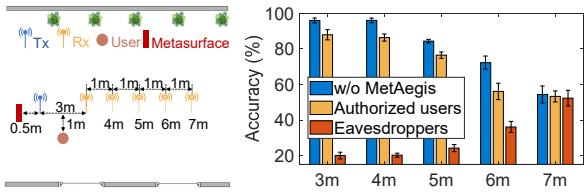
7.3 Evaluation of Extrinsic Impact Factors

To comprehensively evaluate MetAegis's robustness performance, we evaluate a series of external impact factors, including the different environments, different sensing distances, different orientations, different Rx locations, different users, and different transceiver devices.

Impact of different environments. Impact of different environments. We conduct experiments in three different environments: a single apart- ment(representing a complex multipath environment), an office, and a corridor(representing a simple multipath environment). In each environment, we collected data by repeating the five daily activities five times. To verify the MetAegis 's dual capability of protecting legitimate users while resisting eavesdroppers, we considered w/o MetAegis, for authorized users and against eavesdroppers, three personalized configurations. The experiment layouts are shown in Figure 24(a). From the results shown in Figure 24(b), the scenario without MetAegismaintains over 90% accuracy across all environments, demonstrating significant vulnerability to eavesdropping attacks. When MetAegis is applied, the average accuracy for eavesdroppers drops to 19.9%. Crucially, authorized users employing MetAegis's scheme maintain high recognition accuracy (average 88.5%) across all environments. These results conclusively demonstrate MetAegis 's robustness in diverse real-world settings while effectively balancing security and usability.

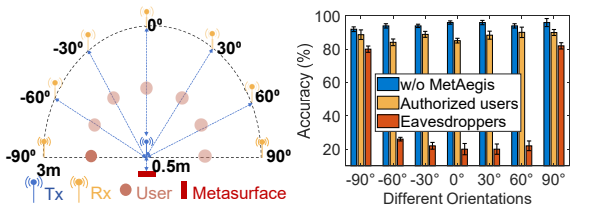
Impact of different sensing distances of transceiver. In order to verify the performance of MetAegis across varying sensing distances, the experiments is conducted with the deployment configuration illustrated in Figure 25(a). We deploy the metasurface at 0.5 m from the transmitter, and the user is fixed at 1 m vertically from the center of the transceiver (when the spacing is 3 m). We move the receiver in steps of 1 m, so that the distance between the transceivers moves from 3m to 7m, and verify the recognition accuracy w/o MetAegis, for authorized users, and against eavesdroppers. As shown in Figure 25(b), two key trends emerge with increasing distances: (1) Both w/o MetAegisand for authorized users scenarios exhibit gradually decreasing accuracy (from 96.1% to 54.4%, and 89.1% to 53.5% respectively). That is because the activity feature is gradually aggravated by multipath interference as the distance increases. (2) Conversely, the recognition accuracy for eavesdroppers is improving (from 20% to 52.3%). That is due to the obfuscation effect of the mask introduced by the metasurface, which is gradually weakening due to the energy attenuation reaching the receiver. Notably, it can be found that the metasurface can provide a 25.1% such low average accuracy for eavesdroppers within a range of 6 m. The observed results verify the effectiveness of obfuscation and the protection performance of MetAegisat different transceiver distances, demonstrating the strong security performance in practical deployment ranges.

Impact of different orientations between metasurface and Rx. In order to verify the defense performance of MetAegisagainst potential eavesdroppers at multiple angles, the experiment is conducted across 180° range (from $-90°$ to $90°$ in steps of $30°$), and the experiment deployment is depicted in Figure 26(a). Specifically, we fix the transmitter at the center of the circle, place the metasurface at 0.5 m from the transmitter, and maintain the distance between the receiver(simulating the eavesdropper) and the transmitter at 3 m. In addition, since the user position may affect the recognition performance of the eavesdropper, assume an ideal scenario where the user always stays within the effective perception link of the receiver regardless of the orientation of the receiver. The experimental results is as shown in Figure 26(b), within $[-60°, 60°]$ range, the eavesdropper's



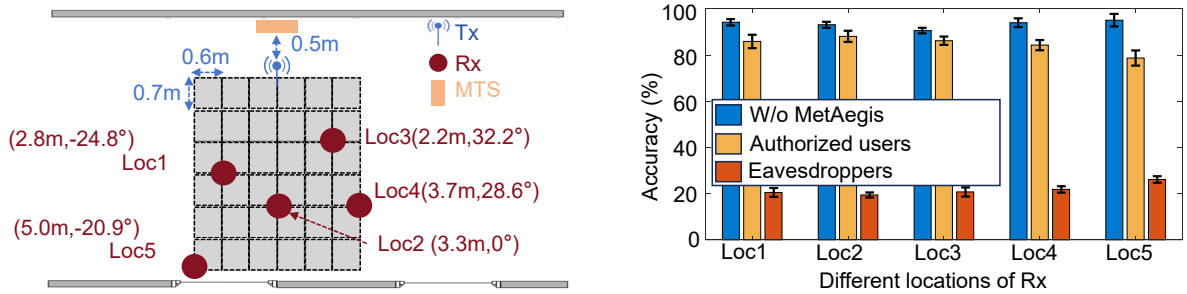
(a) The experiment layout (b) The results of different sensing distances

Fig. 25. The performance of different sensing distances between transceiver.



(a) The experiment layout (b) The results of different orientations

Fig. 26. The performance of preventing different orientations of eavesdroppers.



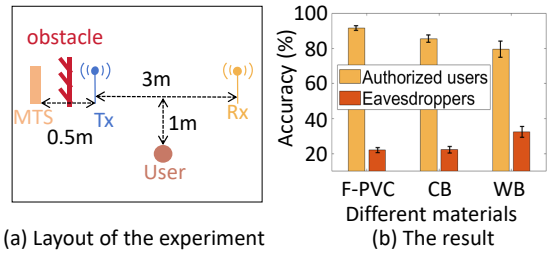
(a) The experiment layout (b) The result of different locations of Rx

Fig. 27. The performance of different locations of receivers.

accuracy is remained below 30% under deploying MetAegiscase. It is worth noting that when the view angle of the metasurface is beyond the boundary (i.e., $\pm 60°$), the obfuscation performance gradually diminished because the metasurface's beam configuration cannot no longer cover the area. However, the system's angular coverage can be expanded by re-adjusting the position/direction of the metasurface or deploying multiple metasurfaces.

Impact of different locations of receivers. To evaluate MetAegis's robustness in real-world scenarios, where the attacker's Rx may deploy in random locations, we conduct rigorous testing under challenging conditions where both the distance and orientations between the MTS and the Rx are all changed. Specifically, to deploy the receiver, sampling five locations within $3.6m \times 4.2m$ rectangular grid area positioned in front of the transmitter. The distance and angle of each location are shown in Figure 27(a). The distance span range is $[2.2m, 5.0m]$ and the angle span range is $[-24.8°, 32.2°]$. The results are shown in Figure 27(b), we find that receiver positioning variations have minimal impact on system performance. The accuracy at Loc5 for eavesdroppers is higher than others, achieving 26.1%, which is due to signal attenuation effects (reduced masking power with increased distance). But it can still maintain 21.7% accuracy on average against eavesdroppers. In addition, the authorized user can achieve 84.6% accuracy for authorized users on average, which demonstrates the robust security protection and authorized access performance across diverse practical deploying scenarios.

Impact of different obstacles. In order to evaluate the robustness, we evaluate the performance when the "Tx-MTS" link is blocked by different daily materials. Specifically, the experiment layout is shown in Figure 28 (a), The obstacles are deployed at the middle of the "Tx-MTS" link. We select three different materials, including foam-PVC board (F-PVC), cardboard (CB), and wood board (WB). The result is shown in Figure 28 (b), the average of the eavesdropping is below 25%, and the average of the legitimate user is above 88%. In addition, it can be found that the accuracy under the wood board reduces below 80%, because the wood board has a larger permittivity and has more serious power attenuation when the signal crosses the obstacles.



(a) Layout of the experiment (b) The result

Fig. 28. The performance of traversing obstacles.

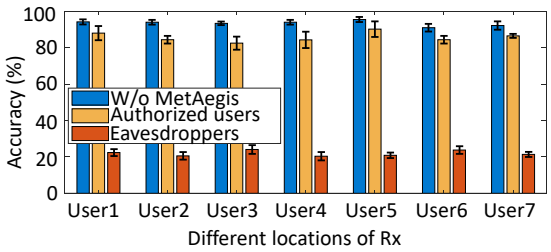
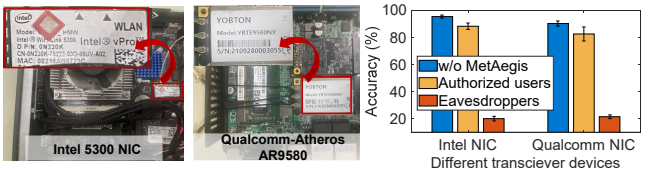
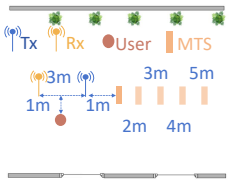


Fig. 29. The performance of different users.

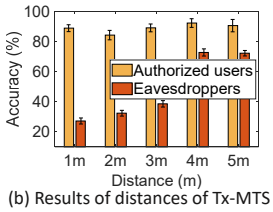


(a) The details of Intel and Qualcomm NIC (b) The result of different transceiver

Fig. 30. The performance of deploying different type transceiver devices.



(a) The experiment layout



(b) Results of distances of Tx-MTS

Fig. 31. The performance of different distances between the metasurface and transmitter.

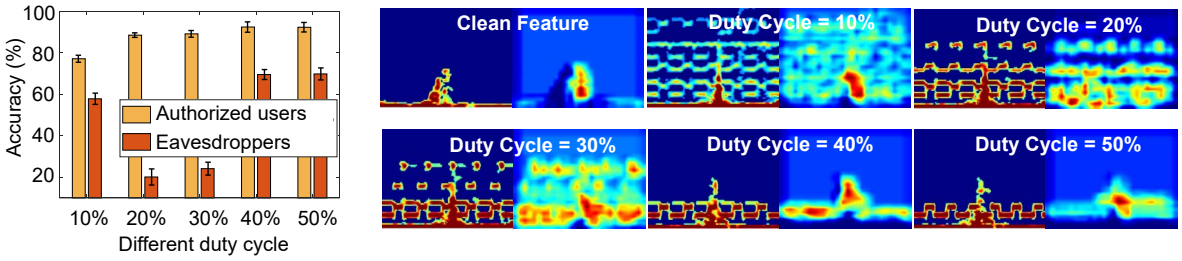
Impact of different users. To evaluate system robustness against user variability, we recruited 7 volunteers to conduct controlled experiments with the deployment, in which the metasurface is 0.5 m behind the transmitter, and the transceivers are 3 m apart. The result is shown in Figure 29. These results validate that MetAegisachieves reliable 85.8% average accuracy for authorized users and 21.9% accuracy on average against eavesdroppers, which demonstrates robust performance across diverse users.

Impact of different transceiver devices for obfuscation and recover. To evaluate system compatibility across different hardware platforms, we conduct experiments with Intel 5300 NIC and Qualcomm-Atheros AR9580 NIC, two widely deployed NICs, which the details are as shown in Figure 30(a). We fix other deployments and change the devices to repeat the experiment. The result is shown in Figure 30(b). The results of accuracy without MetAegisare 95.2% and 90.1% on average, respectively. That observed discrepancy in accuracy likely stems from the hardware error between different devices [59]. The accuracy for authorized users is above 82%, and the accuracy against eavesdroppers is all lower than 21%. It is worth noting that the results of accuracy for authorized users and against eavesdroppers show consistently effective performance.The results demonstrate the effectiveness of MetAegison currently common used transceiver devices.

7.4 Evaluation of Intrinsic Impact Factors

To thoroughly understand the performance of MetAegis, we investigate how the following intrinsic factors affect system effectiveness, including the impact of distance from the sensing links, the impact of duty cycle rate value choice, the impact of long-duration deployment, the impact on the performance of throughput, and the impact under types of disturbance.

Impact of different distances between metasurface and sensing links. Given the potential variations in physical deployment scenarios, the metasurface may be deployed at a different distance from the transmitter. As shown in Figure 31(a), we maintain a fixed transceiver distance of 3m while adjusting the metasurface placement behind the transmitter from 1m to 5m at a step of 1m. The result is shown in Figure 31(b), for the authorized user, the recognition accuracy in different distances is all above 84.1%. The finding confirms the system's resilience even when the metasurface is deployed relatively far away from the transmitter. Besides, we observe an inverse



(a) The result of different duty cycle (b) The result of feature visualization

Fig. 32. The performance of selecting different duty cycle rate values and the comparison by feature visualization.

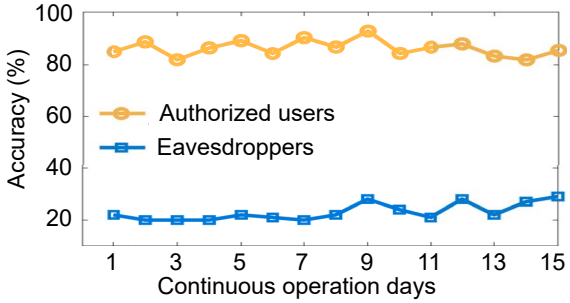


Fig. 33. The performance of continuous work.

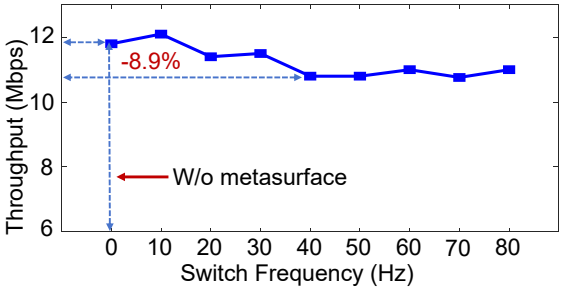
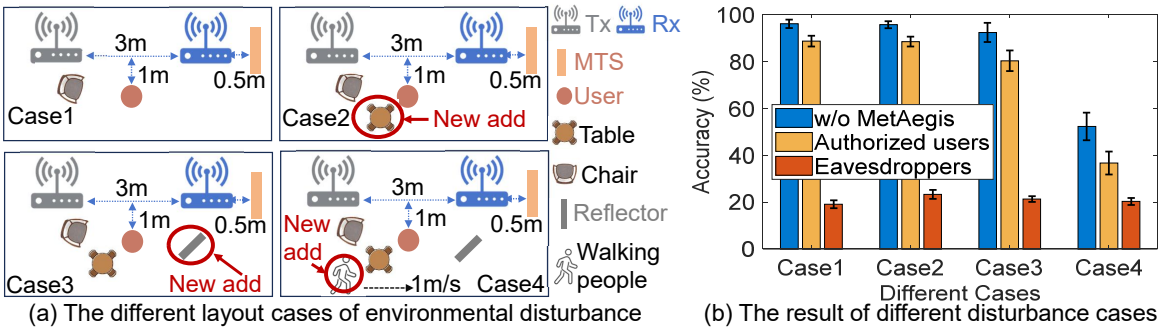


Fig. 34. The performance of data throughput.

relationship between distance and security performance: as the metasurface-transmitter distance increases, the accuracy without recovery shows gradual improvement. In addition, the accuracy against eavesdroppers increases as the distance increases. That is due to the fact that the signal modulated by the metasurface experiences seriously attenuation after long-distance propagation, making it increasingly difficult to effectively mask the original signal features at the receiver. Nevertheless, even at a long distance, the MetAegiscan still drop the accuracy for eavesdroppers above 20% . The metasurface deployment position, which can be flexibly adjusted by users, can be optimally positioned around sensing links to enhance protection.

Impact of different duty cycle value choice. To verify the different duty cycle selected when switching the coding configurations. We maintain the same experimental layout as in previous tests. We vary duty cycle values from 10% to 50% at a step 10%. The result is shown in Figure 32(a), the accuracy reduction rate firstly rises as the duty cycle increases, then drops as it increases. To explore the reason, we employ feature visualization analysis. As shown in Figure 32(b), the power of each frequency component diminishes as the duty cycle decreases, resulting in insufficient feature masking at smaller duty cycles. Conversely, at larger duty cycles, the high-harmonic components will become weaker (as analyzed in Section 5.1), compromising complete the whole feature coverage. That dual-effect creates a “trade-off" duty cycle between power and harmonics. Through comprehensive evaluation, 20% is as the optimal duty cycle that effectively balances those competing factors, and consequently adopt this value for our experimental configurations.

Impact of different deployment times. It’s important to evaluate the long-term effectiveness in preventing eavesdropping. We conducted an extended 15-day deployment experiment, which was operated by pre-saved coding configurations. We keep the metasurface working during the period. The metasurface is positioned at 0.5 m from the transmitter, and the distance of the transceiver is 3 m. Then we repeat the experiments per 24 hours a time. The result is shown in Figure 33, the system consistently maintains an average accuracy for authorized users above 86% while suppressing below 23% accuracy for eavesdroppers, and the worst accuracy during this period for authorized users is still above 80%, which demonstrates persistent security effectiveness. It confirms that MetAegiscan prevent potential eavesdroppers for long-duration deployment.



(a) The different layout cases of environmental disturbance (b) The result of different disturbance cases

Fig. 35. The performance under different types of disturbances (including static and dynamic cases).

Impact on the data throughput of legitimate links. In order to evaluate the impact of the MetAegissystem on the quality of wireless communication links, we conduct throughput measurements using the iperf3 network test tool [6] to build a User Datagram Protocol (UDP) connection, to evaluate the quality of the communication of the metasurface dynamic control channel. Specifically, we establish a performance baseline by conducting 10 repeated measurements in an environment without the metasurface and computing the average throughput. Then, we activate the metasurface and measure throughput at modulation frequencies ranging from 10 Hz to 80 Hz in 10 Hz increments. The experiment result is shown in Figure 34. When the metasurface is turned on, throughput reduction for legitimate links occurs, and the most serious observed degradation is merely 8.9% relative to the baseline, demonstrating that MetAegisdoes not have an extreme impact and maintains minimal impact on the normal communication performance of the legitimate link.

Impact of environmental disturbance. Calling back to the Figure 3(a) shows, typically Wi-Fi sensing results can be formulated as the combination of the static path H_s and dynamic path H_d. Therefore, in order to evaluate the effectiveness under different disturbances in the environment, we consider four typical cases. As shown in Figure 35(a), the first two cases consider the change of the H_s by introducing a chair and table separately into the original environment. For the third case, we introduced a metal reflector to simulate the strong change of H_s effects. Besides, the fourth case introduced dynamic interference of H_d through an unrelated user walking behind the sensing area. As shown in Figure 35(b), the MetAegismaintains stable performance under static disturbances, with authorized users achieving an average accuracy of 85.8%. However, dynamic human movement caused a significant performance degradation. It is due to that the dynamic movement will be added to the objective activity in the feature domain, which will disturb the sensing system's work, and we can see the accuracy without MetAegisdrops to 52.3%. This is because the current sensing model focuses on a single user; the feature of activity will change heavily when other users occur, which will reduce the accuracy of sensing. However, it is worth noting that from the eavesdropper's view, the accuracy remains below 23.3% (almost a guessing probability) in all four cases, demonstrating that MetAegiscan achieve effective and reliable eavesdropping protection regardless of environmental disturbance type.

8 DISCUSSION AND FUTURE WORK

8.1 Generality of MetAegis.

In this part, we discuss the capability of MetAegisbeyond the evaluations we verified. Specifically, we include the following aspects: scenario generalization, sensing system generalization, and sensing task generalization.

Recovery framework generalization. Currently, we use the MAE framework to recover the masked feature. The core principle behind it is related to the key insight of MetAegis, which focuses on removing the mask added by the metasurface in the physical layer using the image reconstruction in the computer vision field. This includes, but is not limited to the MAE framework. Some other framework can also assist MetAegisto achieve this

goal. Specifically, MAGE [25] achieves more advanced performance by changing the input as semantic tokens instead of pixels, which allows the network to operate at a high semantic level without losing low-level details, leading to significantly higher performances than existing methods. Meanwhile, the diffusion model is also a potentially feasible approach to reconstructing the masked feature. DiffIR [55] has shown the performance using the diffusion model to reconstruct high-quality images from low-quality's (even masked images).systemname can be seamlessly compatible with one of these advanced recovery frameworks. In other words, it only needs to replace the current MAE recovery module and may change the input that transfers to the model, which will be our future work.

Sensing system generalization. Using spectrograms or other frequency formats to assist processing is a normal way in wireless sensing [21, 58, 63]. While some systems utilize basic signal characteristics like CSI amplitude and phase as features, it is worth noting that using some advanced signal processing can recover the spectrogram image to the time-domain signal. As shown in Figure 36(a), this is the original spectrogram of the activity, and the time-domain signal of the CSI signal is shown in Figure 36(b). The MAE framework will generate the reconstructed feature as an image output, which is feasible to convert into 2-D matrix data. Here, we use the Inverse Short-Time Fourier Transform (ISTFT) to rebuild the time-domain signal. Although the 2-D matrix data converted from an image loses the phase information. We can use the Griffin-Lim algorithm [43] in the acoustic signal processing field to iteratively optimize the signal phase combination. The results of the rebuilt spectrogram and signal are shown in Figure 36(c-d). It shows the high similarity with the original signal, which can support the following signal processing in the amplitude or phase domain. Therefore, MetAegisdoesn't limit the signal formats. In addition, the obfuscation strategy of MetAegisis based on the introduction of masks in the frequency domain, which fundamentally modulates the amplitude or phase of the CSI signal. Therefore, MetAegiscan be backward compatible with the sensing system using other features in principle. In addition, it should be emphasized that MetAegisfocuses on recovering the activity-based features, and different classifier models can be transparent to the downstream activity recognition implementation, ensuring broad applicability across diverse sensing systems.

Multi-user sensing defending is also a potential future work. Recently, some works have expanded the wireless sensing to multi-user using the modified wide-band transceiver [29], near-field multi-Rx [19], etc. Due to the nature of the MetAegis, which focuses on recovering masked features, it is transparent to the sensing model. Furthermore, as previously demonstrated (Section 6.3), we can achieve multi-directional protection under metasurface FoV coverage, thus achieving encryption for multiple receivers in multi-user sensing. Therefore, MetAegisis theoretically suitable for multi-user sensing protection applications, which will be our future work.

Sensing task generalization. Some extra sensing tasks also carry private information, such as breathing monitoring and indoor tracking, which should also be considered for obfuscation. Here we discuss the feasibility of extending the application of MetAegisto other sensing tasks. Breathing monitoring represents a typically fine-grained sensing task as it relies on precise analysis of CSI amplitude or phase variations. Even if the eavesdropper can monitor information in extreme cases, since the frequency of breathing is very low, a potential method is that can modulate the amplitude or phase of the original CSI signal by quickly changing the coding configuration, thereby obfuscating the original physiological signal. It can confuse the low-frequency private physiological information completely. Similarly, the authorized user can obtain the timing switching sequence of the coding configuration and recover the sensing result. But potential challenge is that the synchronization requirements for the metasurface and the transceiver will become strict, and we will continue to explore it in the future.

For indoor tracking tasks like state-of-the-art work mD-Track [56], it considers using advanced path separation algorithms to iteratively reconstruct the estimated parameters of each multi-path arrival receiver, including time of flight (ToF), angle of arrival (AoA), angle of departure (AoD), etc. Thanks to the dynamic reshaping of the channel by MetAegis, a "virtual path" can be established to combat eavesdroppers. In more detail, the metasurface-generated reflection paths exhibit significantly stronger signal intensity compared to natural multipath components, causing

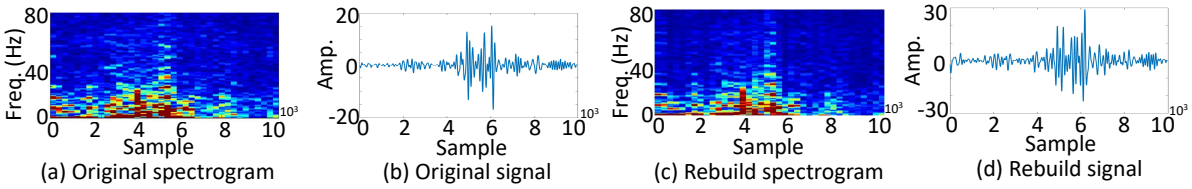


Fig. 36. The processing of the mask token generation from the encrypted feature received by authorized users.

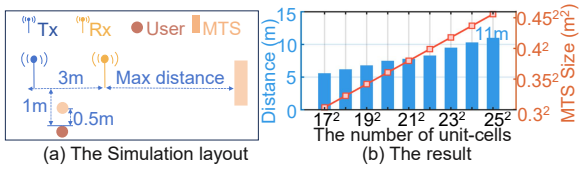


Fig. 37. The Simulation result of protecting distance under different number of unit-cells.

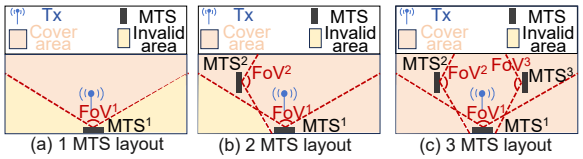


Fig. 38. The case of coverage boundary under different number of metasurfaces.

them to dominate the channel characteristics observed by unauthorized receivers. Thus, it is difficult for the eavesdropper to achieve accurate user tracking. For authorized users, the influence of the "virtual path" of the metasurface can be removed, because of the prior knowledge of the details impact generated by the metasurface. Besides, MetAegiscan generate sidelobe beamforming and introduce controlled phase noise[27]. It will directly affect the eavesdropper's AoA estimation due to the introduction of phase noise.

8.2 Practical Deployment Analyze

In this section, we discuss the practical deployment analysis for the more challenging scenarios and give some feasibility countermeasures.

Protection distance limitations. Since the metasurface only dynamically adjusts the phase of the incident electromagnetic signal and cannot provide additional energy. As the energy of the wireless signal will be severely attenuated as the distance increases, this means that the mask energy that our metasurface can provide is sensitive to distance. Fortunately, more metasurface units can focus more energy on the receiver (as shown in Formula 9). Therefore, increasing the number of metasurface units can increase the protection limit distance of the system. In order to evaluate the limit of the system under different numbers of units, we used Matlab to perform theoretical simulations. Specifically, as shown in Figure 37 (a), we regard the user as a point and the activity as the change in the reflection path distance [62]. Therefore, the change in the reflection path distance can be used to approximately replace the intensity of the user's activity. The intensity of the mask provided by the metasurface can be approximately regarded as the gain of the main lobe pattern under different metasurface units. The power of the metasurface reflected to the receiver can be formulated as [53]:

$$P_r = P_t \frac{G_t G_r (MNd_x d_y)^2}{16\pi^2 (d^t_{m,n} d^r_{m,n})^2} \tag{15}$$

where P_r and P_t are the received signal power and transmit signal power. G_t and G_r are the power of transmitter and the receiver. M and N are the number of cells in each row and column of the square metasurface. d_x and d_y are the length and width of the metasurface. $d^t_{m,n}$ and $d^r_{m,n}$ are the distance of the $(m,n)^{th}$ unit-cell to transmitter and receiver. It is easy to find that the attenuation of pass-loss is only determined by the size of the metasurface and $d^t_{m,n}$ and $d^r_{m,n}$. Besides, the metasurface with more number of unit-cells can provide more signal gain, which can defend the power attenuation. Therefore, We determine that once the power of the mask is weaker than the power of activity, as the distance between metasurface and the transmitter changes, MetAegiswill be invalid

and define the distance as the maximum distance. The result is shown in Figure 37 (b), it can be found that the maximum distance increases as the number of unit-cells increases. When the number of unitcells reach to 25^2, the maximum distance can be up to 11m. Therefore, users can leverage larger metasurfaces to achieve longer protection distances for customized scenarios.

FoV limitations. Due to the limitations of FoV in metasurface, it is hard to cover the full directions to achieve defense against eavesdroppers. Once eavesdroppers move out of the range of FoV, the defense scheme will fail. Nevertheless, the metasurface we used has the range of $(-60°, 60°)$ FoV. Therefore, we can squeeze eavesdroppers in a harsh area that has poor sensing performance due to the extreme sensing angle. In addition, as mentioned in Section 8.1, we emphasize that the user can deploy the metasurface to defend against the high-risk area according to the physical-world sensing scenario. For scenarios where the maximum possible protection angle may be needed, a potential approach is to cascade multiple metasurfaces. In other words, multiple metasurfaces can be deployed to compensate for each other's FoV limitations. A typical example is shown in Figure 38. We sequentially increase the number of deployed metasurfaces with the same FoV. As the number of metasurfaces reaches three, the coordinated FoV covers all angles of the sensing area. Currently, the use of multiple metasurfaces focuses more on achieving coverage of communication blind spots [31, 36]. This will be our future work to introduce multiple metasurfaces to assist sensing.

Deployment on IoT devices. The larger model size and computation demand model like vision transformers limit the deployment on resource-constrained mobile devices. Currently, this has become a common challenge in edge computing and Wi-Fi sensing field. Currently, the parameters of the model after training for reconstruct the feature is 111.6 M with 18.3 GFLOPs, and the memory size is 425.9 MB. The recognition model has 0.65 M parameters with 0.2 GFLOPs. The memory size is 2.5 MB. The time cost for per task in our current laptop (Intel Core i5-8300H) are about 350 ms and 150 ms. A potential solution is that training the recognition model offline and only deploy the final model into IoT devices to achieve wireless sensing. In addition, the existing works have shown the potential that exploiting structured model pruning can save more inference latency and reduce memory by pruning unimportant Attention heads and feed-forward network layers. Meanwhile, efficient vision transformer design is also a new direction to further fit the IoT/edge devices, exciting works like LeViT [14] and MobileViT [39] propose CNN-transformer hybrid models. By combining the strengths of CNN and vision transformer, the hybrid models can achieve comparable accuracy with vision transformers but have a much smaller calculation amount. Therefore, deploying the system on IoT devices like ESP32 [4] to achieve WiFi-aware defense will be our future work.

Defense against advanced eavesdroppers. As for the eavesdroppers who have sufficient activity-related features across different scenarios or an advanced recognition model like a cross-domain Wi-Fi sensing system [60]. It is possible for eavesdroppers to recover the activity-related features by using the existing superior model, such as Generative Adversarial Network (GAN) [20]. However, the model needs sufficient training data, which means the eavesdropper has to collect a large-scale data set for targeted activities and target users in every possible sensing scenario, that will increase the cost of the attack seriously. Face of these advanced eavesdroppers, a potential countermeasure is that MetAegiscan exploit the metasurface and add a "fingerprint" into the feature in the training phase. Therefore, the feature will no longer be repeated by eavesdroppers without any prior knowledge of the fingerprint.

9 CONCLUSION

This paper introduces MetAegis, a metasurface-driven system to protect the privacy of Wi-Fi sensing on the channel side. Previous Wi-Fi sensing protection systems focus on modifying the hardware of sensing devices, which will bring additional hardware or firmware modification costs and can not be extended to deployed devices. Some of them even ignore the normal sensing performance of authorized users. In contrast, MetAegiscan obfuscate

the wireless feature for activities to protect against eavesdroppers by dynamically configuring the coding set on the metasurface near the transmitter. Furthermore, the authorized users can maintain the sensing performance by exploiting the recovery scheme we proposed. Our prototype implementation demonstrates that MetAegiscan reduce the eavesdroppers' recognition accuracy down to 24% on average, and maintain the authorized users' accuracy on average above 88% under the regular deployment. This work represents a paradigm in wireless sensing security by transferring the obfuscation burden from the source to the channel side, overcoming fundamental limitations of traditional source-side protection methods.

Acknowledgments

This work was supported in National Natural Science Foundation of China (Grant number 62276211), Young Scientists Fund of the National Natural Science Foundation of China (Grant Number 62402269), China Postdoctoral Science Foundation (Grant number 2022M722572), and the Humanities and Social Sciences Program of the Ministry of Education (Grant number 24XJCZH021).

References

[1] [n. d.]. Burglaries. https://www.covesmart.com/blog/10-signs-burglars-are-casing-a-house/,.
[2] [n. d.]. CareClaim. https://www.accidentclaimsadvice.org.uk/care-home-negligence-claims/.
[3] [n. d.]. CSI tool. https://dhalperi.github.io/linux-80211n-csitool/,.
[4] [n. d.]. ESP32. https://stevenmhernandez.github.io/ESP32-CSI-Tool/,.
[5] [n. d.]. Home carer stole from elderly. https://www.bbc.com/news/uk-england-kent-19946587,.
[6] [n. d.]. iperf3 network tool. http://software.es.net/iperf,.
[7] [n. d.]. SMP1340-040LF. https://www.skyworksinc.com/Products/ Diodes/SMP1340-Series.,.
[8] Heba Abdelnasser, Khaled A Harras, and Moustafa Youssef. 2015. UbiBreathe: A ubiquitous non-invasive WiFi-based breathing estimator. In *Proceedings of the 16th ACM international symposium on mobile ad hoc networking and computing*. 277–286.
[9] Chen Chen, Yanzhi Ren, Hongbo Liu, Yingying Chen, and Hongwei Li. 2022. Acoustic-sensing-based location semantics identification using smartphones. *IEEE Internet of Things Journal* 9, 20 (2022), 20640–20650.
[10] Yirui Deng, Jiajun Xie, Cong Liu, Xuan Men, Deepak Mishra, and Aruna Seneviratne. 2024. WiFi sensing based textile wetness monitoring. In *ICC 2024-IEEE International Conference on Communications*. IEEE, 4185–4190.
[11] Jianyang Ding, Yong Wang, Hongyan Si, Shang Gao, and Jiwei Xing. 2022. Three-dimensional indoor localization and tracking for mobile target based on WiFi sensing. *IEEE Internet of Things Journal* 9, 21 (2022), 21687–21701.
[12] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. 2020. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929* (2020).
[13] Ruiyang Gao, Wenwei Li, Yaxiong Xie, Enze Yi, Leye Wang, Dan Wu, and Daqing Zhang. 2022. Towards robust gesture recognition by characterizing the sensing quality of WiFi signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 1 (2022), 1–26.
[14] Benjamin Graham, Alaaeldin El-Nouby, Hugo Touvron, Pierre Stock, Armand Joulin, Hervé Jégou, and Matthijs Douze. 2021. Levit: a vision transformer in convnet's clothing for faster inference. In *Proceedings of the IEEE/CVF international conference on computer vision*. 12259–12269.
[15] Yu Gu, Xiang Zhang, Zhi Liu, and Fuji Ren. 2019. Wifi-based real-time breathing and heart rate monitoring during sleep. In *2019 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 1–6.
[16] Kaiming He, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollár, and Ross Girshick. 2022. Masked autoencoders are scalable vision learners. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 16000–16009.
[17] Steven M Hernandez, Deniz Erdag, and Eyuphan Bulut. 2021. Towards dense and scalable soil sensing through low-cost WiFi sensing networks. In *2021 IEEE 46th Conference on Local Computer Networks (LCN)*. IEEE, 549–556.
[18] Jingyang Hu, Hongbo Jiang, Siyu Chen, Qibo Zhang, Zhu Xiao, Daibo Liu, Jiangchuan Liu, and Bo Li. 2024. Wishield: Privacy against wi-fi human tracking. *IEEE Journal on Selected Areas in Communications* (2024).
[19] Jingzhi Hu, Tianyue Zheng, Zhe Chen, Hongbo Wang, and Jun Luo. 2023. MUSE-Fi: Contactless muti-person sensing exploiting near-field Wi-Fi channel variation. In *Proceedings of the 29th annual international conference on mobile computing and networking*. 1–15.
[20] Shady Abu Hussein, Tom Tirer, and Raja Giryes. 2020. Image-adaptive GAN based reconstruction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34. 3121–3129.

[21] Sijie Ji, Yaxiong Xie, and Mo Li. 2022. SiFall: Practical online fall detection with RF sensing. In *Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems*. 563–577.
[22] Chenghan Jiang, Jinjiang Yang, Xinyi Li, Qi Li, Xinyu Zhang, and Ju Ren. 2024. RISiren: Wireless Sensing System Attacks via Metasurface. In *Proceedings of the 2024 on ACM SIGSAC Conference on Computer and Communications Security*. 3332–3345.
[23] Chenggao Li, Qianyi Huang, Yuxuan Zhou, Yandao Huang, Qingyong Hu, Huangxun Chen, and Qian Zhang. 2023. Riscan: Ris-aided multi-user indoor localization using cots wi-fi. In *Proceedings of the 21st ACM Conference on Embedded Networked Sensor Systems*. 445–458.
[24] Chenning Li, Manni Liu, and Zhichao Cao. 2020. WiHF: Gesture and user recognition with WiFi. *IEEE Transactions on Mobile Computing* 21, 2 (2020), 757–768.
[25] Tianhong Li, Huiwen Chang, Shlok Mishra, Han Zhang, Dina Katabi, and Dilip Krishnan. 2023. Mage: Masked generative encoder to unify representation learning and image synthesis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2142–2152.
[26] Tianxing Li, Xi Xiong, Yifei Xie, George Hito, Xing-Dong Yang, and Xia Zhou. 2017. Reconstructing hand poses using visible light. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 1–20.
[27] Xinyi Li, Chao Feng, Fengyi Song, Chenghan Jiang, Yangfan Zhang, Ke Li, Xinyu Zhang, and Xiaojiang Chen. 2022. Protego: securing wireless communication via programmable metasurface. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking*. 55–68.
[28] Xinyi Li, Chao Feng, Xiaojing Wang, Yangfan Zhang, Yaxiong Xie, and Xiaojiang Chen. 2023. {RF-Bouncer}: A Programmable Dual-band Metasurface for Sub-6 Wireless Networks. In *20th USENIX Symposium on Networked Systems Design and Implementation (NSDI 23)*. 389–404.
[29] Xin Li, Hongbo Wang, Zhe Chen, Zhiping Jiang, and Jun Luo. 2024. Uwb-fi: Pushing wi-fi towards ultra-wideband for fine-granularity sensing. In *Proceedings of the 22nd Annual International Conference on Mobile Systems, Applications and Services*. 42–55.
[30] Xinyu Li, Jian Wei You, Ze Gu, Qian Ma, Long Chen, Jingyuan Zhang, Shi Jin, and Tie Jun Cui. 2024. Passive human sensing enhanced by reconfigurable intelligent surface: Opportunities and challenges. *IEEE Communications Magazine* (2024).
[31] Xinyi Li, Gaoteng Zhao, Ling Chen, Xinyu Zhang, and Ju Ren. 2024. RFMagus: Programming the Radio Environment With Networked Metasurfaces. In *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*. 16–30.
[32] Fen Liu, Jing Liu, Yuqing Yin, Wenhan Wang, Donghai Hu, Pengpeng Chen, and Qiang Niu. 2020. Survey on WiFi-based indoor positioning techniques. *IET communications* 14, 9 (2020), 1372–1383.
[33] Haipeng Liu, Kening Cui, Kaiyuan Hu, Yuheng Wang, Anfu Zhou, Liang Liu, and Huadong Ma. 2022. mTransSee: Enabling environment-independent mmWave sensing based gesture recognition via transfer learning. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 1 (2022), 1–28.
[34] Hao Liu, Hanting Ye, Jie Yang, and Qing Wang. 2021. Through-screen visible light sensing empowered by embedded deep learning. In *Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems*. 478–484.
[35] Jun Luo, Hangcheng Cao, Hongbo Jiang, Yanbing Yang, and Zhe Chen. 2024. MIMOCrypt: Multi-user privacy-preserving Wi-Fi sensing via MIMO encryption. In *2024 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2812–2830.
[36] Ruichun Ma, Shicheng Zheng, Hao Pan, Lili Qiu, Xingyu Chen, Liangyu Liu, Yihong Liu, Wenjun Hu, and Ju Ren. 2024. Automs: Automated service for mmwave coverage optimization using low-cost metasurfaces. In *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*. 62–76.
[37] Yongsen Ma, Gang Zhou, and Shuangquan Wang. 2019. WiFi sensing with channel state information: A survey. *ACM Computing Surveys (CSUR)* 52, 3 (2019), 1–36.
[38] Chunxu Mao, Mohsen Khalily, Pei Xiao, Long Zhang, and Rahim Tafazolli. 2020. High-gain phased array antenna with endfire radiation for 26 GHz wide-beam-scanning applications. *IEEE Transactions on Antennas and Propagation* 69, 5 (2020), 3015–3020.
[39] Sachin Mehta and Mohammad Rastegari. [n. d.]. MobileViT: Light-weight, General-purpose, and Mobile-friendly Vision Transformer. In *International Conference on Learning Representations*.
[40] Mark J Nave. 1989. The effect of duty cycle on SMPS common mode emissions: theory and experiment. In *Proceedings, Fourth Annual IEEE Applied Power Electronics Conference and Exposition*. IEEE, 3–12.
[41] Sameera Palipana, David Rojas, Piyush Agrawal, and Dirk Pesch. 2018. FallDeFi: Ubiquitous fall detection using commodity Wi-Fi devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 4 (2018), 1–25.
[42] Chaoyi Pan, Dan Xu, Shiyuan Xu, Fuwei Wang, Chao Zheng, Bingyue Liu, and Xiaojiang Chen. 2024. Perfat: Non-contact Abdominal Obesity Assessment System. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 4 (2024), 1–25.
[43] Nathanaël Perraudin, Peter Balazs, and Peter L Søndergaard. 2013. A fast Griffin-Lim algorithm. In *2013 IEEE workshop on applications of signal processing to audio and acoustics*. IEEE, 1–4.
[44] Yue Qiao, Ouyang Zhang, Wenjie Zhou, Kannan Srinivasan, and Anish Arora. 2016. {PhyCloak}: Obfuscating sensing from communication signals. In *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16)*. 685–699.

[45] Sridhar Rajagopal. 2012. Beam broadening for phased antenna arrays using multi-beam subarrays. In *2012 IEEE International Conference on Communications (ICC)*. IEEE, 3637–3642.

[46] Tina Samajdar and Md Iqbal Quraishi. 2015. Analysis and evaluation of image quality metrics. In *Information Systems Design and Intelligent Applications: Proceedings of Second International Conference INDIA 2015, Volume 2*. Springer, 369–378.

[47] Kalpana Seshadrinathan and Alan C Bovik. 2008. Unifying analysis of full reference image quality assessment. In *2008 15th IEEE International Conference on Image Processing*. IEEE, 1200–1203.

[48] Zhambyl Shaikhanov, Sherif Badran, Hichem Guerboukha, Josep M Jornet, Daniel M Mittleman, and Edward W Knightly. 2024. MetaFly: Wireless Backhaul Interception via Aerial Wavefront Manipulation. In *2024 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2759–2774.

[49] Jayanth Shenoy, Zikun Liu, Bill Tao, Zachary Kabelac, and Deepak Vasisht. 2022. Rf-protect: privacy against device-free human tracking. In *Proceedings of the ACM SIGCOMM 2022 Conference*. 588–600.

[50] Eric A Silva, Karen Panetta, and Sos S Agaian. 2007. Quantifying image similarity using measure of enhancement by entropy. In *Mobile multimedia/image processing for military and security applications 2007*, Vol. 6579. SPIE, 219–230.

[51] Paul Staat, Simon Mulzer, Stefan Roth, Veelasha Moonsamy, Markus Heinrichs, Rainer Kronberger, Aydin Sezgin, and Christof Paar. 2022. IRShield: A countermeasure against adversarial physical-layer wireless sensing. In *2022 IEEE Symposium on Security and Privacy (SP)*. IEEE, 1705–1721.

[52] Yuqi Su, Fusang Zhang, Kai Niu, Tianben Wang, Beihong Jin, Zhi Wang, Yalan Jiang, Daqing Zhang, Lili Qiu, and Jie Xiong. 2024. Embracing distributed acoustic sensing in car cabin for children presence detection. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 1 (2024), 1–28.

[53] Wankai Tang, Xiangyu Chen, Ming Zheng Chen, Jun Yan Dai, Yu Han, Marco Di Renzo, Shi Jin, Qiang Cheng, and Tie Jun Cui. 2022. Path loss modeling and measurements for reconfigurable intelligent surfaces in the millimeter-wave frequency band. *IEEE Transactions on Communications* 70, 9 (2022), 6259–6276.

[54] Wei Wang, Alex X Liu, Muhammad Shahzad, Kang Ling, and Sanglu Lu. 2017. Device-free human activity recognition using commercial WiFi devices. *IEEE Journal on Selected Areas in Communications* 35, 5 (2017), 1118–1131.

[55] Bin Xia, Yulun Zhang, Shiyin Wang, Yitong Wang, Xinglong Wu, Yapeng Tian, Wenming Yang, and Luc Van Gool. 2023. Diffir: Efficient diffusion model for image restoration. In *Proceedings of the IEEE/CVF international conference on computer vision*. 13095–13105.

[56] Yaxiong Xie, Jie Xiong, Mo Li, and Kyle Jamieson. 2019. MD-Track: Leveraging Multi-Dimensionality for Passive Indoor Wi-Fi Tracking. In *International Conference on Mobile Computing and Networking (MobiCom)*.

[57] Zheng Yang, Yi Zhang, Guoxuan Chi, and Guidong Zhang. 2022. Hands-on wireless sensing with Wi-Fi: A tutorial. *arXiv preprint arXiv:2206.09532* (2022).

[58] Zheng Yang, Yi Zhang, Kun Qian, and Chenshu Wu. 2023. {SLNet}: A spectrogram learning neural network for deep wireless sensing. In *20th USENIX Symposium on Networked Systems Design and Implementation (NSDI 23)*. 1221–1236.

[59] Enze Yi, Fusang Zhang, Jie Xiong, Kai Niu, Zhiyun Yao, and Daqing Zhang. 2024. Enabling WiFi sensing on new-generation WiFi cards. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 4 (2024), 1–26.

[60] Guolin Yin, Junqing Zhang, Guanxiong Shen, and Yingying Chen. 2022. FewSense, towards a scalable and cross-domain Wi-Fi sensing system using few-shot learning. *IEEE Transactions on Mobile Computing* 23, 1 (2022), 453–468.

[61] Youwei Zeng, Dan Wu, Jie Xiong, Jinyi Liu, Zhaopeng Liu, and Daqing Zhang. 2020. MultiSense: Enabling multi-person respiration sensing with commodity WiFi. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 3 (2020), 1–29.

[62] Youwei Zeng, Dan Wu, Jie Xiong, Enze Yi, Ruiyang Gao, and Daqing Zhang. 2019. FarSense: Pushing the range limit of WiFi-based respiration sensing with CSI ratio of two antennas. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 1–26.

[63] Jin Zhang, Bo Wei, Wen Hu, and Salil S Kanhere. 2016. Wifi-id: Human identification using wifi signal. In *2016 international conference on distributed computing in sensor systems (DCOSS)*. IEEE, 75–82.

[64] Jia Zhang, Rui Xi, Yuan He, Yimiao Sun, Xiuzhen Guo, Weiguo Wang, Xin Na, Yunhao Liu, Zhenguo Shi, and Tao Gu. 2023. A survey of mmWave-based human sensing: Technology, platforms and applications. *IEEE Communications Surveys & Tutorials* 25, 4 (2023), 2052–2087.

[65] Xianan Zhang, Yu Zhang, Guanghua Liu, and Tao Jiang. 2023. AutoLoc: Toward ubiquitous AoA-based indoor localization using commodity WiFi. *IEEE Transactions on Vehicular Technology* 72, 6 (2023), 8049–8060.

[66] Yangfan Zhang, Yaxiong Xie, Zhihao Hui, Hao Jia, and Xiaojiang Chen. 2024. Hydra: Attacking OFDM-base Communication System via Metasurfaces Generated Frequency Harmonics. In *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*. 938–952.

[67] Yi Zhang, Yue Zheng, Kun Qian, Guidong Zhang, Yunhao Liu, Chenshu Wu, and Zheng Yang. 2021. Widar3. 0: Zero-effort cross-domain gesture recognition with Wi-Fi. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44, 11 (2021), 8671–8688.

[68] Yue Zheng, Yi Zhang, Kun Qian, Guidong Zhang, Yunhao Liu, Chenshu Wu, and Zheng Yang. 2019. Zero-effort cross-domain gesture recognition with Wi-Fi. In *Proceedings of the 17th annual international conference on mobile systems, applications, and services*. 313–325.

[69] Guozhen Zhu, Beibei Wang, Weihang Gao, Yuqian Hu, Chenshu Wu, and KJ Ray Liu. 2024. WiFi-based robust human and non-human motion recognition with deep learning. In *2024 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*. IEEE, 769–774.